DIFFERENTIAL PREDICTIVE VALIDITY OF HIGH SCHOOL GPA AND COLLEGE ENTRANCE TEST SCORES FOR UNIVERSITY STUDENTS IN YEMEN

by

Abdulghani Ali Dawod Al-Hattami

B.A., Hodeidah University, 1998

M.Ed., Hodeidah University, 2000

M.A., Loyola University Chicago, 2006

Submitted to the Graduate Faculty of

the School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Abdulghani Ali Dawod Al-Hattami

It was defended on

February 24, 2012

and approved by

Suzanne Lane, Ph.D., School of Education

Louis A. Pingel, Ph.D., School of Education

Lauren Terhorst, Ph.D., School of Nursing

Dissertation Advisor: Clement A. Stone, Ph.D., School of Education

DIFFERENTIAL PREDICTIVE VALIDITY OF HIGH SCHOOL GPA AND COLLEGE ENTRANCE TEST SCORES FOR UNIVERSITY STUDENTS IN YEMEN

Abdulghani Ali Dawod Al-Hattami, PhD

University of Pittsburgh, 2012

High school grade point average and college entrance test scores are two admission criteria that are currently used by most colleges in Yemen to select their prospective students. Given their widespread use, it is important to investigate their predictive validity to ensure the accuracy of the admission decisions in these institutions.

This study was conducted to investigate the predictive validity of both high school GPA and college entrance test scores used as predictors in the admission process to Yemeni colleges. In addition, the differential predictive validity of high school GPA and college entrance test scores was examined across gender and high school location. The relationship between students' persistence in the four years of college and the predictor variables, high school GPA and college entrance test scores, was studied as well. College entrance test scores were examined for their reliability in order to evaluate any potential influence of measurement error on the analyses.

The sample in the study consisted of 881 cohort students from two public universities in Yemen. The data analysis for this study included reliability coefficient analysis, multiple and logistic regression analyses, and Gulliksen and Wilks (1950) tests for differential prediction. Results showed that high school GPA and college entrance test scores were both significant predictors of academic performance as measured by first-year college GPA and four-year

cumulative GPA. However, differential predictive evidence was observed for different gender and rural/urban subpopulations. The results implied that using common regression equations to predict academic performance may result in unfair admission decisions. Finally, while high school GPA was not a significant predictor of college persistence, the predictive validity was enhanced when college entrance test scores were added to the prediction equation of college persistence. In summary, high school GPA explained a very small portion of the total variance of first-year college GPA and four-year cumulative GPA. Therefore, a comprehensive review of the use of high school GPA for admission decisions is strongly recommended.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

My sincere thanks go to the Fulbright Program, sponsored by the United States Department of State, and the Research Methodology Program at the School of Education at the University of Pittsburgh for the financial assistance, without their help I would not have been able to afford getting my M.A. and Ph.D. in the United States.

In the Research Methodology program at the University of Pittsburgh, many people deserve my thanks. First, I would like to express my deepest appreciation and thanks to Dr. Clement Stone, my advisor and dissertation chair. There is no doubt in my mind that his knowledge, experience and patience were the bridge that helped me to overcome the obstacles that I encountered while I was working to achieve my goals. His friendly treatment and conversations gave me the confidence that I needed to complete this degree. To Drs. Suzanne Lane, Louis Pingel, Kevin Kim, and Feifei Ye, thank you for your excellent teaching and instruction. Thank you for helping me improve my knowledge and skills in educational measurement, statistics, and research design.

I would also like to thank my dissertation committee members, Professor Suzanne Lane, Professor Louis Pingel, and Dr. Lauren Terhorst. They deserve my deepest appreciation for their guidance, comments, suggestions and time.

Special thanks to Dr. Abdu Hodeish, Vice President for student affairs at Hodeidah University and Dr. Abdulsalam Al-Iryani, Dean of student affairs at Ibb University, for providing me with the data and information needed to make this research possible.

My heartfelt gratitude goes to my parents, brothers and sisters in Yemen for their prayers, encouragement and emotional supports. Also, my deepest appreciation and thanks for my friends and any person, who helped and supported me directly or indirectly throughout the journey of my study, for their encouragement to complete my higher education.

Above all, I could not have accomplished my Ph.D. degree without the big love of my family. Absolutely there are no words that can sufficiently express the deepest love and appreciation that I have for them. I deeply appreciate the love and inspiration that you have so generously given to me. My wife, Entesar, you made the family complete while I was absent for years. It was your support, understanding, patience, and confidence that inspired and helped me complete this journey of Ph.D. program. You have stood by me through the many difficult situations that I faced; therefore, I sincerely thank you.

Finally, to my children, Ranyah, Rana, Aseel, and Ruba, thank you so much. Your love and smiles were the main source of my energy and dedication to accomplish my goal. Thank you my family, you understood why I was unable to spend as much family time with you as we supposed to do.

1.0 INTRODUCTION

When the number of applicants exceeds the capacity of postsecondary institutions, it must be decided which students are more qualified and most likely to succeed in these institutions. Selection criteria vary from one institution to another and from one country to another, and deciding which criteria are most accurate in predicting academic success in postsecondary institutions is a complex task. Cognitive factors (e.g., SAT scores), noncognitive factors (e.g., personality traits), and demographic characteristics (e.g., gender, ethnicity, location) are major criteria for the admission decisions in most of the postsecondary institutions around the world. Nevertheless, future academic success has, traditionally, been predicted from cognitive factors used as the sole criteria of academic success (Pentages & Creedon, 1978).

Literature shows that both high school grade point averages and standardized test scores, such as the SAT or ACT, are generally significant predictors of student success during their undergraduate studies (Astin, Korn, & Green, 1987; Noble, 1991; Moffat, 1993; Bridgeman, McCamley-Jenkins, & Ervin, 2000; Snyder, Hackett, Stewart, & Smith, 2003; Kim, 2002; Kuncel, Hezlett, & Ones, 2004; Ramist, Lewis, & McCamley-Jenkins, 1994; Waugh, Micceri, & Takalkar, 1994; Wolfe & Johnson, 1995; Kuncel et al., 2005; Kuncel, Credé, & Thomas, 2007). Moreover, a significant body of literature suggests that high school GPA more accurately predicts academic success in college than standardized tests or any other factor (Munro, 1981;

Lawlor, Richman, & Richman, 1997; Peltier, Laden, & Martranga, 1999; Snyder, Hackett, Stewart, & Smith, 2003; Camara & Echternacht, 2000; Tross et al, 2000; Fleming & Garcia, 1998; Fleming, 2002; Hoffman, 2002; Zheng et al., 2002; Gose, 1994). However, other studies on the predictive validity of entrance test scores (such as the SAT) suggest that the prediction validity of academic success is enhanced by using entrance test scores (Camara & Echternacht, 2000).

In countries other than the U.S., for example in Yemen, postsecondary institutions select promising applicants based only on their levels of performance on high school grade point average. Engineering, medical, and science colleges use entrance test scores in addition to high school GPA. Using these indicators to make important decisions about the students' future academic career raises the question about their validity in predicting future academic success. Many students are denied entrance into particular programs as a result of insufficient test scores. The question of whether they are sufficient predictors of future academic success should be investigated in order to ensure fair admission decisions.

However, using high school GPA as a criterion for admission into postsecondary institutions has been criticized by educators, parents, and faculty at most Yemeni universities for various reasons. First, high school GPA is not a cumulative measurement of the entire high school career of a student but is calculated based on student scores gained in the twelfth grade only. Second, tests are centralized. They are constructed by the Ministry of Education and administered to all twelfth grade students once a year, after students have completed the school year. Some schools or teachers might not have covered the specified topics due to many reasons (e.g., getting textbooks weeks after schools have started, a lack of teachers for some science subjects in many rural areas). Third, in Yemen, there are two majors available to students in

secondary school: the Science track and the Literary track. These majors have slight differences in regard to their standards, curricula, and number of subjects. Fourth, high school general exams are developed to measure what students have been taught during the twelfth grade. This results in teachers sometimes "coaching for the test." They direct their teaching solely to preparation for the exam, devoting more time to exposing their students to the previous forms of high school graduation exams than other material. They also provide students with summaries and answers to expected questions on the test to be memorized. Fifth, there are considerable differences between schools located in rural and urban areas. Schools in rural areas lack appropriate equipment, qualified teachers, and supervision. There has been no consideration of this discrepancy during the admission process.

Finally, there is a common pervasive phenomenon that might affect the use of high school GPA: cheating. Due to the prevalence of corruption and the power of personal influence, many relatives try to help their children cheat on the exams in various ways (e.g., by bribing the proctors or superintendents or using their high positions to enable their children to cheat). While the admission to higher education institutions is becoming more competitive, it is becoming possible that a one-point difference on a test may affect the likelihood of a student being admitted. Furthermore, given that high school GPA is dependent upon the chosen major, high school location, and students' backgrounds, the validity of using only high school GPA (at all humanities colleges) or both high school GPA and college entrance test scores (at all applied colleges) for admission decisions has become a concern.

College entrance tests have also been criticized as they are not different from high school exams, that is, they are still tests that measure the students' knowledge of subjects taken in high school. The college entrance test is composed of a combination of some of the following subject areas: English, Physics, Chemistry, Biology, and Mathematics. Each college determines what combination of subjects would help them select their prospective students. For students who meet the specified high school GPA required by the college, the final admission decision is based on the overall weight of both high school GPA and college entrance test scores. The disadvantages associated with using high school GPA and college entrance test scores make these factors inaccurate predictors of students' abilities and achievements required in higher education.

1.1 STATEMENT OF THE PROBLEM

Higher education in Yemen has witnessed a dramatic expansion in the last two decades with enrollment increasing approximately 16% each year since 1990 (Education Indicators in the Republic of Yemen, 2009). In 1990, there were only two universities. By 2001, the number had increased to include 15 universities (seven public and eight private). At that time, the government of Yemen decided that the minimum required percentage score for admission into all postsecondary institutions should be a high school GPA of 70% or higher (Education Indicators in the Republic of Yemen, 2009). Colleges of Science, Medicine, and Engineering and some other departments require college entrance test scores in addition to high school GPA to determine the admission of a student. Precise entry requirements may vary from year to year.

These measures are used in postsecondary institutions in Yemen not because of their validity or their predictive power of college success but to limit the number of applicants to match the capacity of the colleges. It is also assumed that by raising the admission criteria, more qualified students will attend the colleges.

Considering that high school GPA and college entrance test scores are very important criteria for admission to higher education and therefore have serious consequences for students, it is crucial to investigate whether high school GPA and college entrance test scores accurately predict the future academic success of students at universities in Yemen. Exhaustive research on this topic has not produced any study specifically focused on the predictive validity or the psychometric quality of high school GPA and/or college entrance test scores. This study was conducted to investigate reliability and validity evidence for using high school GPA and college entrance test scores for admission decisions by institutions of higher education in Yemen. The effects of gender and high school location on the predictive validity of high school GPA and college entrance test scores were examined.

1.2 RESEARCH QUESTIONS AND HYPOTHESES

After a review of the literature regarding the prediction of college student academic performance, five research questions were formed. These questions are:

1. What is the reliability, as measured by internal consistency, of the college entrance test scores used for admission into Yemeni universities?

2. Are high school GPA and college entrance test scores significant predictors of first-year college GPA for each of the different colleges at Yemeni universities? Does the addition of college entrance test scores enhance the prediction of college performance? Are the results comparable to what has been found in the United States?

Hypotheses: High school GPA and college entrance test scores make significant contributions to the prediction of first-year college GPA for each of the different colleges at Yemeni universities, and the addition of college entrance test scores enhances the prediction of college performance. The results are hypothesized to be comparable to what has been found in the United States.

3. Do high school GPA and college entrance test scores have differential prediction when used to predict academic performance in Yemeni universities across gender and high school location? Are the results comparable to what has been found in the United States?

Hypotheses: High school GPA and college entrance test scores have differential prediction when used to predict academic performance in Yemeni universities across gender and high school location. Female students are expected to outperform male students and urban school students perform better than rural school students.

4. How well do high school GPA and college entrance test scores predict students' long-term academic success in college as measured by four-year cumulative GPA? Are the results comparable to what has been found in the United States?

Hypotheses: High school GPA and college entrance test scores make a significant contribution when used to predict academic success over time. The results are expected to be comparable to what has been found in the United States.

5. How well do high school GPA and college entrance test scores predict students' persistence to graduation?

Hypotheses: High school GPA and entrance test scores make a significant contribution in predicting students' persistence to graduation.

1.3 SIGNIFICANCE OF THE STUDY

The findings of this study can be valuable in three ways. First, they may guide admissions personnel and decision-makers at the Ministry of Higher Education and Scientific Research in identifying whether high school GPA and college entrance test scores are accurate predictors of academic performance of students attending higher education institutions. It might help them in the development of future admission plans and student retention programs at Yemeni universities and colleges. Further, the results of this study can help high school counselors at the Ministry of Education assist with the college transition needs of their graduating students, by being able to better identify students at risk for dropping out. Second, the findings might guide the educational stakeholders in Yemen to review the testing policy as well as the quality of high school assessments and college entrance tests. Third, this study will bridge a research gap in the study of academic performance of students attending postsecondary institutions in Yemen and thus serves as a motivation for future research to be conducted in this area.

1.4 DEFINITION OF TERMS

For the purpose of this study, a glossary of commonly used terms and concepts is necessary:

Academic performance: a reference to how well a student performs in academic knowledge and skills which is reflected by that student's cumulative grade point average (GPA).

Correlation coefficient: a statistical index of the linear relationship between two variables or measures. Coefficients range from –1.00 to +1.00 with values near zero indicating no relationship and values far from zero indicating a strong relationship; positive correlations mean that high values on both variables occur jointly while negative correlations mean an inverse relationship exists between the variables. In test validity studies, correlation coefficients between a predictor and a criterion are often called validity coefficients.

Criterion: an outcome or dependent variable or test score. In predictive validity studies, the criterion most frequently used is the first-year college grade point average. Other criteria used include cumulative college GPA and retention.

College Entrance Test: a test used to assess the student's readiness for admission into a graduate institution.

Differential prediction: refers to a situation where the best prediction equations and/or the standard errors of estimate are significantly different for different groups of examinees.

Differential validity: refers to a situation where the computed validity coefficients are significantly different for different groups of examinees.

High school GPA: a percentage score that is calculated based on the total weighted scores obtained from the high school general examination results of all subjects studied during the 12th grade only.

High school major: a major in high school that a student can choose to study. Students can focus on a specific area in each major. These majors are the *Science* and *Literary* tracks.

Imam: an Islamic word that refers to a leader of a mosque or community.

Persistence: refers to whether or not a student continues in college until he/she graduates from the same institution in a norm-time of four years. Persistence is synonymous with retention (Braxton, 2003).

Prediction equation: the resulting equation obtained from a linear regression analysis with a single criterion and one or more predictors computed from a sample of students.

Predictive validity: one of the aspects of test validity as originally defined by the American Psychological Association. It is most commonly used to describe the relationship between a predictor such as a test score and a criterion such as a grade point average.

Predictor: an independent variable used to forecast a criterion variable. In predictive validity studies, the most commonly used predictors are one or more test scores and high school grade point average. Typically, the predictor scores are temporally available before the criterion scores.

Retention: refers to a program outcome in which a student continues to attend the same institution and eventually earns a degree or certificate of graduation (Giordano, 1996).

1.5 ORGANIZATION OF THE STUDY

This study is organized into five chapters. Chapter one provides the reader with background information about the study and introduces the statement of the problem, the research questions, and the significance of the study. Chapter two provides some background on the Yemeni education system. In addition, it sheds some light on the history of the educational system before and after the unification of north and south Yemen in 1990. Studies on standardized admission tests and their prediction power are reviewed in this chapter. The chapter also discusses the reliability and validity of test scores and how test bias could be detected. It ends with a review of retention studies. Chapter three describes the overall methodology: the design of the study, the data source, the sample, the data collection procedures, and the statistical analysis procedures employed. Chapter four provides the findings of the study corresponding to each research question. The final chapter presents a summary and interpretation of the research findings along with the conclusion and implications of the major findings, limitations of the present study, and recommendations for future research.

2.0 LITERATURE REVIEW

Identifying reliable and valid predictors of students' future academic success is a concern to educators, researchers, and admission committees. It has been a topic of many investigations over the years. This chapter provides a review of existing literature related to this issue. First, the chapter presents a background of the Yemeni education system and an overview of predictions of college academic success. The chapter also provides an overview of the differential validity and differential prediction studies. It ends with a discussion of the importance of retention studies.

Since there are not many studies relevant to the issue in question in the Republic of Yemen, the main part of previous research is built on admission measures used in the United States. Numerous reliability and validity studies have focused on the United States admission system, but the findings do have valuable potential implications for admission systems elsewhere in the world.

2.1 BACKGROUND OF THE YEMENI EDUCATION

2.1.1 General Information about Yemen

The Republic of Yemen is situated in the southwestern corner of the Arabian Peninsula. It is bordered by the Kingdom of Saudi Arabia to the north, the Gulf of Aden and Arabian Sea to the south, the Red Sea to the west, and the Sultanate of Oman to the east (Figure 1). It is approximately 532,000 square kilometers in size. Sana'a is the capital of the country. The nation is divided into 21 administrative governorates. The population in Yemen is estimated to be 23 million people. Arabic is its official language. Additionally, Yemen is the only country in the Arabian Peninsula to have a republican form of government.

Figure 1. Map of Yemen (From: http://www.semp.us/publications/biot_reader.php?BiotIDD=680)

Yemen is categorized as the least developed country in the Middle East with a GDP per capita of US$631. According to a UNICEF report, more than 40% of the country's population is considered below the poverty line, and about 16% of the population has PPP (Purchasing Power Parity) of less than $1 a day (Human Development Report, UNDP, 2005). The unemployment rate in Yemen is approximately 40% as of 2007 (Hill, 2007). Yemen has some relatively small oil reserves that were discovered during the 1980s, though they are expected to be depleted by 2017, which may lead to an economic collapse (Fontaine and Exum, January 5, 2010). However, Yemen has large proven reserves of natural gas (Encyclopedia Britannica, 2010). The first liquefied natural gas production launched in October 2009.

2.1.2 Educational System in Yemen

The educational system in Yemen can be categorized in two periods: the first pertains to the pre-unification period, in what was called North Yemen and South Yemen, and the second is the post-unification period (after May 22, 1990). This section shows how the two education systems merged into one and how the higher education system has developed.

2.1.3 Before the Unification – North Yemen and South Yemen

2.1.3.1 North Yemen (1919-1990)

During the rule of Imam Yahya from 1919 until 1948, North Yemen was an isolated country with very few connections to the outside world. This isolation resulted in a poor education system. Imam Yahya built some religious schools during his reign; these schools were usually attached to mosques, and students were taught by Islamic scholars. Imam Yahya isolated the country from the outside world as a political strategy to protect his rule until his assassination in 1948. His son, Ahmad, succeeded him.

Imam Ahmad (1948-1962) introduced a modern education system as part of his attempts to open Yemen to the outside world. More secular schools were established late in his reign, but the number of schools was not sufficient for the population. During that time, many Yemenis went to Egypt, Lebanon, or Europe for better education. Imam Ahmad died in 1962 and was succeeded by his son Mohammed Al-Badr; however, a revolt headed by opposition groups followed soon after they deposed the Imam.

After the September 22, 1962 revolution, a new republican government came to power. During that period, many secular and vocational schools were established. New subjects like English, mathematics, and social and natural sciences were introduced. Furthermore, schools for girls were established for the first time in some major cities. Greater educational and intellectual awareness spread, public education grew, and libraries were established. In addition, the first Ministry of Education was established in 1963 to monitor the public school system (MoE, 2010).

In the 1970s and 1980s there was a tremendous improvement and expansion in the educational system that made it more accessible. In 1981, there was about a 50% attendance rate for students compared to 10% in 1971. However, the number of female students did not improve significantly, remaining between 10-15%. In addition, the retention rate for primary school was low at about 12% in the late 1970s. From 1962 to 1990, although there was remarkable improvement, it did not meet the expectations of the Yemeni people, especially when compared to education in neighboring countries. During that time, the school system consisted of six years of primary, three years of preparatory, and three years of secondary education.

The most important step in the development of higher education in Yemen was the establishment of Sana'a University in 1970. Sana'a University was the only university in the Yemen Arab Republic (YAR) until the unification with the People's Democratic Republic of Yemen (PDRY) in 1990. Sana'a University began with three colleges, the College of Law, College of Science, and College of Arts. In the late 1970s, some other colleges were added, including the College of Economics and College of Education (See Appendix C for a complete list of all colleges and their years of establishment).

2.1.3.2 South Yemen (1839-1990)

The British ruled South Yemen from 1839 until 1967. During the British occupation, there was limited access to education. In the 1930s there were very few public and private schools; however, after World War II, there was a slight improvement in the structure of the public education system. In the 1940s the British introduced English lessons to be taught in schools. Furthermore, some successful young people were sent to the United Kingdom for higher

education. Like the YAR, there were growing political awareness and intellectual movements during the 1950s and early 1960s. Intellectuals and cultural elite formed civil groups and encouraged the use of the Arabic language in everyday life. After the independence of the country in 1967, the education system of south Yemen was left in a precarious situation because it had relied on foreign teachers during the British occupation (Dresch, 2000).

After independence, the education system in the People's Democratic Republic of Yemen (PDRY) expanded considerably and was nationalized, while access to it was greatly broadened. The public education system was similar in basic structure to that of the YAR, with six years of primary, three years of preparatory, and three years of secondary level. In 1975 this structure was changed to two years of preschool, eight years of primary education, and four years of secondary education. Vocational programs and some specialized secondary programs were available as well. Although education was not obligatory, it was completely free at all levels. Furthermore, free textbooks and transportation were provided to all children, even those in rural areas. A monthly stipend was given to students at the university level to encourage higher education enrollment (MoE, 2010). However, the female population was still underrepresented; the primary school enrollment was at about 20-25% in the 1970s, growing to approximately 35-40% in the 1980s (MoE, Statistical Year Book, 2006/2007).

The University of Aden was established in 1970, as the only university in the People's Democratic Republic of Yemen, with six colleges. These colleges were for education, law, economics, agriculture, technology, and medicine. Some teaching institutes were also built to produce graduates who could teach immediately at primary schools. The College of Education at the University of Aden provided teachers who were able to teach at the secondary level.

2.1.4 The Republic of Yemen (1990)

At the time of the unification of the YAR and PDRY in May 1990 and the establishment of the Republic of Yemen, the two countries had poor education systems and faced the same challenges. They shared the same sources of problems, including a lack of financial resources, a poor infrastructure that disadvantaged rural areas, and limited teacher skills. After unification, the texts and curricula were standardized, and the education system was slightly restructured to include nine years of primary education and three years of secondary education. Access to education for rural areas and female students remained a substantial problem, and the dropout rate of students at the primary school level was still high.

The Constitution of the Republic of Yemen that was enacted in 1994 specified the role of the government in terms of education in Article 53:

> Education is a right for all citizens. The state shall guarantee education in accordance with the law through building various schools and cultural and educational institutions. Basic education is obligatory. The state shall do its best to obliterate illiteracy and give special care to expanding technical and vocational education. The state shall give special attention to young people and protect them against perversions, provide them with religious, mental and physical education, and the appropriate environment to develop their aptitude in all fields.

The Yemeni government provides a universal, compulsory, and free education for children ages 6 through 15 (General Law for Education, 1992). Although this requires the government to supply fair and equitable access to education for all Yemeni children, results have been disappointing in tracking illiteracy and equal education rates for males and females. For

example, although the government is required to provide an equal opportunity to all students regardless of their race or ethnicity, there are some restrictions in terms of gender. Some majors are only open to males, such as physical education and some technical majors at postsecondary institutions.

The educational system in Yemen is supervised by three departments: the Ministry of Education, which oversees the public education of both the government and private sector, the Ministry of Higher Education and Scientific Research, which oversees higher education in the government and private sectors and research centers, and the Ministry of Technical Education & Vocational Training, which oversees the technical education and industrial, agricultural, commercial, and technical vocational training. The focus of this research is on public and higher education.

2.1.4.1 Public Education System

The Yemeni public education system is operated and supervised by the Ministry of Education. The Ministry sets standards, develops curricula, and designs textbooks for all schools in the country. The education system in Yemen includes nine compulsory years of basic education followed by three years of general secondary education. Preschools and kindergartens are not widely implemented; however, there are a growing number of public and private ones, particularly in major cities. Students must be six years old to be admitted to primary education (MoE, 2010; National Information System, 2010).

The academic year at the two stages (primary and secondary education) consists of two semesters with approximately 15 weeks each. Public schools for boys and girls are entirely separated from each other, with schools for boys operated by men only and those for girls run by women only. However, private schools are both coeducational and single-sex.

2.1.4.2 Secondary (High) School

General secondary education covers three years. In the first year, students follow the same curriculum: Holy Quran, Islamic Education, Arabic, English, Mathematics, Physics, Chemistry, Biology, History, Geography, and Society. After completing the first year, students can choose to follow either the science or literary track for the remaining two years. Students in the science track study Mathematics, Physics, Chemistry, and Biology; those in the literary track study History, Geography, Sociology, Economics, Statistics, Psychology, Philosophy, and Logic. In addition, all students in both tracks study Holy Quran, Islamic Education, Arabic, and English (Appendix B). It is worth noting that there are some other schools students can attend beyond the primary school. Those schools include technical and industrial education, agricultural, industrial, commercial, and vocational training as well as some medical and military schools. Those schools are not operated by the Ministry of Education (National Information System, 2010).

The academic year runs from early September to the end of June. To pass a grade, a student has to achieve a minimum percentage score of 50% in each subject. If a student fails to score a percentage score of 50% in any subject during a particular grade, he/she has one chance to retake the test before the beginning of the next academic year. If the student fails to pass, he/she must repeat the same grade. At the end of each academic year, successful students are

awarded a certificate and are promoted to the next grade. At the end of the third year, students must pass the General Secondary Examination with a minimum percentage score of 50% in each subject and a minimum overall percentage score of 70% to be able to pursue studies at postsecondary institutions (MoE, 2010).

2.1.4.3 Higher Education in Yemen

Yemen has seen a rapid growth in higher education since the 1990s, from two universities at the time of unification to 15 universities (seven public and eight private) in the mid-1990s. Today, there are eight public universities (with the addition of Amran University in 2005), each having many different colleges and programs (Appendix C). The eight public universities in Yemen include 105 colleges that are divided into 45 applied colleges and 60 humanities colleges.

The number of students enrolled in the Yemeni universities increased from about 35,000 students in 1990 to reach 233,903 students (including private universities) in 2007, more than six times the original number in seven years. The enrollment of females in university education increased from 16% in 1990 to 28.5% in 2007. The enrollment of students in private universities also increased from almost 0% of the total enrolled in university education in 1993 to about 17.5% in 2007. In 2005/2006, there were about 174,000 students in public universities and about 12,000 in private ones. In 2006/2007, the number of enrolled students in public universities reached 188,557 students (Appendix D) with 45,560 students in private universities in the same year.

The graduates are the final product of university education and are considered an important asset to the development process of the country. The total number of graduates for the year 2007/2008 was 21,697 compared to 16,125 in 2002/2003 (Appendix D). There are also

some institutions that offer two-year degrees (e.g., Technical and Vocational Training Centers); however, enrollments in these institutions are low. The private sector is small, accounting for about 2% of basic and secondary education and 15% of university enrollments in 2004/2005 (Education Indicators in the Republic of Yemen, 2009).

2.1.5 College Admission System in Yemen

In the 1970s and 1980s, there was a seat for every student who wanted to continue higher education at one of the postsecondary institutions. The policy at that time was intended to prepare as many Yemeni citizens as possible to replace the non-Yemeni employees. Therefore, the institutions focused on the quantity rather than quality of students, and they admitted almost any high school graduate that applied.

After 1990, the admission process in postsecondary institutions began to differ from one college to another with a minimum requirement of a high school GPA of 60%. The number of students coming to the university has drastically increased since that time to reach about 176,000 students in 2003 compared to about 35,000 students in 1990 (Higher Education National Strategy, 2009). In 2001, when the Yemeni government noticed the number of prospective students exceeded the number of available places at the universities, it specified that the minimum requirement for admission into all postsecondary institutions should be a high school GPA of 70%. In addition, the government decided that the application process in the colleges of science, medicine, dentistry, and pharmacology should combine college entrance test scores with high school GPA as a way to reduce the number of applicants and receive more qualified ones. It recommended that these colleges should develop their own standardized entrance exam (MoHE,

2010). This entrance exam was intended to give an indication about future academic success at those particular colleges, as government officials believed that it would be insufficient to admit applicants to science colleges based only on high school GPA.

More specifically, the Ministry of Higher Education and Scientific Research stipulated that students who wish to join any college in the Yemeni universities must abide by the following:

a. The applicant must hold the Secondary General Certificate provided that for Yemenis: not more than five years must have passed since their obtainment of the certificate and for others: not more than two years should have passed since their obtainment of the certificate.
b. The applicant must have obtained the percentage score required for admission to the desired college as determined by the council yearly.
c. The applicant must meet all conditions declared in the year he/she is applying and must submit all the required documents which should be in original form.
d. The applicant can apply for two colleges or two majors in one college he/she wishes to join as a first choice and as a second choice provided he/she can meet the requirements.
e. Applicants must sit for the entrance examination in the colleges that require it. Admission is implemented by comparing applicants, taking into consideration 50% of the high school GPA and 50% of the college entrance test scores.

Differences in admission to postsecondary institutions are related to high school major, high school GPA, and college entrance test scores. As mentioned in the secondary school section, there are two majors students can choose to study after successfully completing the 10th grade. Students can choose either the science track or the literary track. This choice is very important because it could decide the student's future career. Students graduating from the science track can be admitted to any college as far as their high school GPA permits. On the other hand, students graduating from the literary track can *only* be admitted to humanities colleges.

While all colleges require a certain high school GPA, science colleges require students to take an entrance test after provisionally accepting all of the students that have the required high school GPA. The college entrance test is then conducted, and the best students are admitted according to the capacity of the college or major. Since it is difficult to account for the total variance that contributes to college success, the focus of this study is on the cognitive factors, high school GPA and college entrance test scores, as predictors of future academic success. Willingham (1974) indicated that high school GPA is a principal criterion that is used in the admissions process because it is readily available, quantifiable, equitable, and assumed to be fair. High school GPA is also assumed to measure desired behaviors like intelligence, aptitude, and achievement that are required for students' subsequent studies (Gottheil & Michael, 1957; Hirschberg, 1977; Humphreys, 1962). Ebel (1978) stated that the basic rationale for using standardized entrance test scores, with high school GPA, in the admissions process is that they are relevant and reliable. They provide admissions committees with a standardized measure of academic achievement for all examinees (Ebel, & Frisbie, 1991).

2.1.6 Gender and Geographical Disparities in Yemeni Education

Gender and geographical disparities in access to education are pronounced in Yemen and might have an impact on students' high school GPA. Schools in large urban areas like Sana'a and Aden are different from schools in small towns or villages in terms of size, accessible resources, qualified teachers, and supervision. Therefore, the geographical location of high schools might affect students' performance which can in turn negatively or positively impact their high school GPAs. Woodruff & Ziomek (2004) reported that "A particular student's grades depend not only on the student's achievement, but also on the school the student attends."

In Yemen, boys and urban children enjoy greater educational opportunities and higher enrollment rates than others. While about 70% of the population lives in rural areas, the basic education net enrollment rate is only 48%, compared to 80% in urban areas. The same scenario is true for secondary education (MoE, Illiteracy and Adult Education Strategy, 2008). The gender inequities are among the largest in the world. According to the 1997/1998 Education Census, female students' basic and secondary net enrollment rates are only 38% and 10%, compared to 71% and 23% for male students. Moreover, the primary school retention rate for female students is only 33%, compared to 68% for male students. Enrollment rates have achieved an average annual growth of 2.04% during the period 2003/2004-2007/2008, with a variation between male students (0.59%) and female students (4.35%). Enrollment has increased from 4,544,746 students in the 2003/2004 academic year to reach to 4,750,588 in the 2007/2008 with a difference of 205,842, which signifies an average growth of 41,168 enrolled students every year (Appendix D).

It has been noted that low female participation in secondary schools is rooted in female retention in basic education. Overall, about 25% of female students complete ninth grade, therefore limiting the number of girls entering secondary education. Both economic factors (e.g., unavailability of schools that are close to home, difficulty in going to schools that are far away, shortage of female teachers, lack of well-equipped schools, etc.) and cultural factors (e.g., social factors, poverty, early marriage, etc.) act as constraints that play an important role in limiting secondary school participation and therefore higher education (Education Indicators in the Republic of Yemen, 2009).

Since the 1990s public education has continuously expanded, which has led to the growth of university students' enrollment during the past period. These numbers have increased from 181,350 students in 2002/2003 to 199,268 students in 2007/2008, which is a figure well above the capacity of the universities (Appendix D). Thus, university education has seen a growth in the number of students from 181,350 students in 2002/2003, of which 25.6% were female students, to 199,268 students, of which 30.5% were female students, in the year 2007/2008. The figures show low female participation despite their increasing numbers during the mentioned period.

In terms of high school performance, Abbass (2011) found that female students outperform male students in both science and literary sections of high school. He also noted that students who come from well maintained (qualified teachers, good regulatory system, proper facilities, etc.) schools tend to perform better than others. Literature, however, revealed no study that has examined the students' performance on college entrance tests in Yemeni colleges.

2.2 PREDICTION OF FUTURE ACADEMIC SUCCESS

Prediction of future academic success falls within the realm of predictive validity evidence. Predictive validity evidence indicates how well an assessment can predict scores obtained at a later time through the use of either the same measure or a different measure. In the *Standards for Educational and Psychological Testing*, predictive validity is defined as "how accurately test data can predict criterion scores that are obtained at a later time" (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999, p. 180). Predictive validity comes into play when a test is used to predict the likelihood of some future performance. It indicates the extent to which an individual's future level on the criterion is predicted from prior test performance (Messick, 1989; Crocker & Algina, 1986).

No research has been conducted on predictors of future academic success in the Republic of Yemen. However, numerous studies have investigated the best predictors of future academic success in U.S. postsecondary institutions. Despite these studies, there is no complete explanation of the variance of academic success in these institutions (Nettles, 1991; Sackettet al., 2001). Admissions personnel need some standards on which to base their admission decisions. They have traditionally relied on cognitive predictors such as high school GPA and standardized test scores to differentiate between applicants.

Literature shows that research on future academic success has discussed two types of predictor variables, namely cognitive and noncognitive predictors. Cognitive predictors refer to the standardized entrance tests such as the SAT and ACT tests and achievement measures such as high school GPA and high school rank-in-class. Noncognitive predictors refer to two main

attributes: personality characteristics, such as self-motivation, self-directedness, dedication to studies, and social skills; and environment factors, such as size of schools, location of schools, parental education, and socioeconomic status (Klugh & Bierly, 1959; Misanchuk, 1977; Himelstein, 1965; Wolfe & Johnson, 1995; Johnson, 2002; Mulvenon, Stegman, Ganley, & McKenzie, 2002; Barnett, Ritter, & Lucas, 2003).

Much research has been conducted to determine which would be more accurate predictors of future academic success in postsecondary institutions. Some researchers favor cognitive predictors (Noble, 1991; Baird, 1984; Bridgeman, McCamley-Jenkins, & Ervin, 2000; Kuncel et al., 2005; Kuncel, Nisbet, Ruble, & Schurr, 1982; Kuncel, Hezlett, & Ones, 2001, 2004; Kuncel, Credé, & Thomas, 2007), whereas others lean toward using noncognitive variables and claim that they are important for the prediction of students' academic success (Pentages & Creedon, 1978; Duran, 1986; Tracey & Sedlacek, 1984; Sedlacek, 2004). Micceri (2001) indicated that cognitive predictors like high school GPA and standardized test scores are the best predictors with variables such as race, ethnicity, and gender providing some additional information. Alderman (1999) also showed that high school GPA is a better predictor of future academic success than other factors such as the demographic variables of race, gender, or socioeconomic status.

The vast majority of the predictive validity research used cognitive predictors such as high school GPA and standardized test scores to predict college academic success (Odell, 1927; Mathiasen, 1984; Ramist, Lewis, and McCamley-Jenkins, 1994; Travers, 1994; Camara & Echternacht, 2000; Hu, 2002; Willingham, 1985; Kuncel et al., 2001, 2004; Bridgeman, McCamley-Jenkins, & Ervin, 2000; Kuncel & Hezlett, 2007; Kuncel, Credé, & Thomas, 2007). Mathiasen (1984), for example, in a review of more than 60 studies, posited that high school

GPA and standardized entrance test scores are the best predictors of college success as they account for approximately 25% of the variance when predicting first-year college GPA. A meta-analysis study by Elert (1992) identified that high school grades were twice as good a predictor of college success than college entrance test scores, with college entrance test scores accounting for approximately 5% of additional variance in the prediction model. Elert noted that the strength of college entrance exams is in the prediction of first-year college grade point averages, and that their predictive power disappears after the first year.

The American Association of College Registrars and Admissions Officers survey ranks high school GPA and rank-in-class as the best predictors of college success. The *Trends in College Admission 2000* reported that between 1979 and 2000, either high school GPA or rank-in-class was consistently the primary factor in admissions decisions (Breland, Maxey, Gernand, Cumming, & Trapani, 2002). Ramist, Lewis, & McCamley-Jenkins (1994) conducted a comprehensive study of the predictive validity of high school GPA and the SAT scores as predictors of college performance. They analyzed data from a total of about 45,000 students from 45 colleges. The results of using only high school GPA as a predictor of first-year college GPA yielded a moderate correlation of .39. When SAT scores were added to the prediction equation, the correlation increased to .48. Their results, like many other studies, showed that high school GPA predicts better than SAT scores, but adding SAT scores to the model increases the validity coefficient by almost .10 beyond high school performance. This indicates that using both high school GPA and standardized test scores improves the accuracy of prediction. Nevertheless, many other studies showed that the correlation of SAT scores or ACT scores with first-year college GPA is about .40, on average. However, using standardized test (ACT or SAT) scores as predictors along with high school GPA yields correlations with first-year college GPA that

average about .50 (The College Board and ETS, 1998; Camara & Echternacht, 2000; Willingham, 1998; Ramist et al., 1994; ACT, 1997; Noble, 1991). In a recent study, Sackett, Kuncel, Arneson, Cooper, & Waters (2009) also examined different studies and found a moderate correlation between standardized admission tests and academic performance (r = .44). The relationships were strong even after controlling for noncognitive factors like socioeconomic status.

Burton & Ramist (2001), in a report summarizing the results of 19 studies based on 227 institutions and over 64,000 students of the relationship between SAT and cumulative college GPA upon completing college, also found that SAT had correlations averaging about .40 with the cumulative college GPA. They also reviewed the relationship between standardized test scores and college graduation. They concluded that "there is a solid academic component to graduation that is measured by [SAT scores and high school record]" (p. 17).

Obviously a large number of predictive validity studies show that high school GPA is consistently the best predictor of first-year college GPA and that standardized test scores do add a statistically significant increment to the prediction. Hence, using both high school GPA and standardized test scores makes better predictions than using high school GPA alone. High school GPA does account for a large amount of the variation in first-year college GPA. Perfetto (2002) stated, "The combination of high school grades and standardized test scores has been part and parcel of evaluating applicants for admissions" (p. 31). Literature, however, showed some limitation of these traditional predictors to significantly predict college success and persistence (Tracey & Sedlacek, 1984; Levin & Wyckoff, 1994; Burton & Ramist, 2001). Opponents of traditional predictors argued that these tests measure only examinees' cognitive skills and that they do not offer a fair prediction of college performance for students from minority groups.

Those researchers encouraged the use of noncognitive predictors (Wigdor & Garner, 1982; Tinto, 1993; Tracey & Sedlacek, 1984).

Past studies have also examined the role of noncognitive predictors of academic success such as metacognitive skills (e.g., Zeegers, 2001), study attitudes (e.g., Zimmerman, Parks, Gray, & Michael, 1977), study motivation (e.g., Melancon, 2002), and even personality traits (e.g., Ridgell & Lounsbury, 2004). Credé and Kuncel (2008) in a meta-analysis study examined the predictive validity of the study habits, skills, and attitude inventories and constructs in predicting academic performance. They found that study habits, skills, and study motivation among other attitudinal constructs accounted for incremental variance (ranging from .04 to .12) in academic performance beyond high school GPA and standardized tests.

Willingham (1985) examined 33 related predictor variables of college success. He categorized them into five groups: traditional academic predictors, supplementary achievement measures, admissions staff rating, goals and plans, and background characteristics. The traditional academic predictors, such as high school grade point average or standardized test scores, are referred to as cognitive measures, and supplementary achievement measures, such as honors, community service, athletics, leadership, creative achievement, follow-through, work experience, personal statement, and school preference are referred to as noncognitive measures. He concluded that high school GPA was the best predictor of future academic success followed by standardized test scores. He found that the validity coefficient for the standardized test scores and high school grades was .43 with first-year college GPA. Other noncognitive measures related to college academic success include parental education, anxiety levels, study skills, motivation, and attitudes to name a few (Mulvenon, Stegman, & Thorn, 1999). Results from research that combines cognitive and noncognitive measures find that noncognitive measures

add little increment to the prediction of successful college performance (Schmitt, Keeney, Oswald, Pleskac, Quinn, Sinha, & Zorzie, 2009).

There are a number of potential benefits to be gained from broadening the selection criteria beyond traditional cognitive predictors, such as the SAT or ACT and high school GPA, but one important benefit is the potential increase in the diversity of students admitted into colleges. Whereas minority students often score lower on cognitive ability tests such as the SAT and ACT, there are small or no differences between majority and minority groups on many noncognitive assessments of background, interests, and motivation (Hough, 1998; Sackett, Schmitt, Ellingson, & Kabin, 2001). A challenge, however, in including these noncognitive predictors and broadening the selection criteria is how to maintain an objective means of comparing applicants on the basis of not only their cognitive ability but also their noncognitive measures.

Since most predictive validity studies have concluded that high school GPA and standardized entrance test scores were more strongly related to college performance than any other predictors, this study focused on these two cognitive predictors. High school GPA and standardized entrance test scores could be used to make accurate predictions and appropriate admission decisions (Eimers & Pike, 1997; Willingham, 1985; Mouw & Khanna, 1993). Noble (1991) emphasized that the logic behind this conclusion is that the cognitive predictors, such as high school GPA and standardized test scores, are related directly to academic skills required for success in college courses.

Research suggests that any measure used to make important educational decisions about students should be reliable (Salvia & Ysseldyke, 2007); this criterion was examined in this study to estimate the reliability coefficient of college entrance test scores used by postsecondary institutions in Yemen. Although there are three main recognized methods for estimating the reliability coefficients (see McDonald, 1999), this study used only the internal consistency approach using Cronbach's alpha (Cronbach, 1951) to obtain the reliability coefficients for all subtests within the college entrance tests and the stratified alpha coefficient (Cronbach, Schoneman, & McKie, 1965) for the composite test at each college. Estimating internal consistency reliability was explained as "dividing the test into two or more constituent parts and in some way estimating reliability from the consistency of performance across these part-tests" (Haertel, 2006, p. 71). A number of internal consistency reliability estimates have been proposed, but the most widely used is Cronbach's coefficient alpha (Haertel, 2006). Cronbach's alpha helps to measure the homogeneity or heterogeneity of the given scales (by finding the variances of all scores for each item and the sum of these variances across all items). Researchers have concluded that the more the items measure the same behavior, the higher the consistency (Anastasi and Urbina, 1997).

Haertel (2006) claimed that the use of the coefficient alpha has several theoretical advantages. First, since it equals the mean of all possible split-half reliability coefficients (i.e., another estimate of internal consistency reliability that involves the division of the total test into two "parallel" sub-tests), the use of the coefficient alpha avoids the arbitrary choice of a split or division. Second, it is mathematically equivalent to one of the lower bounds of the theoretical reliability coefficient. The implication of this is that the theoretical reliability coefficient is higher than the observed coefficient alpha.

For the composite tests, stratified alpha coefficient was used. Stratified alpha was proposed by Cronbach, Schoeneman, and McKie (1965). It is intended for cases where components of a test can be grouped into subtests on the basis of content. It assumes *k* components, where *i*th component (*i* = 1, …. , *k*) consisted of n_i components. Stratified alpha is obtained by

$$Stratified\ \alpha = 1 - \frac{\sum_{i=1}^{k} \sigma_i^2 (1-\alpha_i)}{\sigma_x^2} ,$$

where σ_i^2 is the variance of items in the *i*th component, α_i is the reliability of *i*th component, and σ_x^2 is the variance of the test. Kamata, Turhan, and Darandari (2003) concluded that stratified alpha was the most "reliable" procedure compared to two alternative methods (Maximal Reliability and Multidimensional omega).

2.3 TEST BIAS

One important problem a researcher may face when conducting a predictive validity study is test bias. Test bias is of great social and educational importance and has received much attention over the years (Cole, 1972). While some researchers have made inferences to the fact that admission tests help assess competence of examinees irrespective of their group membership (Jensen, 1980), others have disagreed with such statements. For example, Crouse & Trusheim (1988) mentioned that some critics have stated that college-bound students in different subgroups have different standardized test scores, which usually favor whites and men. In the same vein, Crouse and Trusheim declared that colleges, especially selective colleges that insist that all examinees should have "equal" test scores, end up rejecting minorities and women who typically have

lower scores than the majority of men, thus leading to the under-representation of minorities and women in student populations. Similarly, Nettles and Nettles (1999) stated that bias in standardized tests may create admission barriers to higher education for minorities and women. As a result, many researchers have argued that the use of the standardized tests to screen applicants for admission to colleges is biased against women and minorities. Based on such remarks, it would appear that this issue is very relevant to university admissions in Yemen, as women in Yemen are considered a minority due to many restrictions imposed by society and culture.

Numerous studies have, therefore, been directed to define and identify test bias (House & Keeley, 1993). Cleary (1968) stated that "A test is biased for members of the subgroup of the population if, in prediction of a criterion for which the test was designed, consistent nonzero errors of predictions are made for members of the subgroup" (p. 115). Other researchers have provided different definitions. For example, Shepard (1987) defined bias using one word, "invalidity." Shepard implied that bias brings about systematic errors in the validity evidence gathered for a test. He explained that a systematic error in a test distorts the meaning of the measurement for the members of a particular group. Jensen (1980) also defined bias as "systematic measurement error related to the use of a test with two or more specified populations" (p. 328). Linn (1984) defined bias as a systematic tendency for a test to over- or under-estimate the true abilities of members of a group of examinees classified by demographic variables such as gender and ethnicity. Camilli and Shepard (1994) stated that test bias refers to "invalidity or systematic error in how a test measures for members of a particular group" (p. 8).

In summary, bias arises when test scores result in different meanings for the different identifiable subgroups. The fact that different groups have different average scores does not necessarily imply that the test is biased. Although large between-group differences may be indicative of test bias, score differences are not in themselves sufficient to judge whether a test is biased or not (Wightman, 2003). Test bias is a concept that is defined in terms of groups of examinees. Most often test bias is an issue in the study of gender and ethnic group differences (Camilli & Shepard, 1994, p. 8). To date, there is not a single study that has been conducted on gender-related bias in admission to postsecondary institutions in Yemen. In this study, the differential predictive validity of high school GPA and college entrance test scores across gender and high school location was investigated.

2.3.1 Test Bias based on Gender and High School Location

The literature on this topic consistently reveals that males and females differ in their performance on various high school subjects and standardized tests. In a review of past research on gender differences in test performance, Wilder and Powell (1989) surveyed studies that addressed undergraduate, graduate, and professional school entrance tests, validity studies, national studies, verbal ability tests, and quantitative ability tests. Specific testing programs discussed in the studies reviewed included the National Assessment of Educational Progress, National Longitudinal Study of 1972, High School and Beyond, and the SAT. They found that females outperformed males on verbal ability and achievement tests while males outperformed females on mathematics tests. Although the study revealed that disparities existed between males and females, it also mentioned that these disparities were diminishing slowly over time.

Willingham and Cole (1997), in a comprehensive examination of gender differences, found that gender differences occur across different testing programs and in different subject areas. According to their findings, females tend to achieve better grades in school while males tend to receive better scores on standardized tests. Although some researchers have reported contradictory findings, the results regarding specific tests have generally shown that males tend do better in mathematics and science-related subjects while females perform better on verbal subjects (Azen, Bronner, & Gafni, 2002). Hyde, Fennema, and Lamon (1990) conducted a meta-analysis study and showed that while girls tend to do slightly better in mathematics compared to boys in elementary and middle school, this disparity switches in high school and college, with males tending to do much better than females.

Maccoby and Jacklin (1974) reported that females always have an advantage over males on verbal subjects, and the difference has become more pronounced over time. On the contrary, Hyde and Linn (1988) analyzed 165 studies covering a very large number of participants that reported data on gender differences in verbal ability. Their results showed only a slight difference in females' superiority with the difference being so small that it appeared to be nonexistent. Further, an examination of gender difference by age revealed countering results to the conclusions drawn by Maccoby and Jacklin (1974). That is, their analysis showed no changes in the magnitude of gender differences at different levels or grades.

Research on school location has produced mixed results. Some believe that education in urban schools is better than education in rural schools while others disagree. Young (1998), for example, found that "location of the school ha[s] a significant effect upon student achievement, with students attending rural schools not performing as well as students from urban schools" (p.386). However, according to Young, all the differences were much more related to

socioeconomic status (SES) than location of the school. Like Young, Geske, Grinfelds, Dedze, and Zhang (2006) speculated that the urban/rural disparities in learning achievement can be attributed, at least partially, to differences in the socioeconomic conditions. Bylund & Reeves (2005), on the other hand, stated that "recent research does not provide clear evidence that rural schools are inferior to urban schools". He also concluded that "analysis reveals that rural schools achieve mean annual gains in performance that equal or better their urban counterparts".

A comparison of the performance of urban and rural school students on standardized entrance tests has not produced definitive results. Several studies have not found any significant differences between urban and rural students. Monk and Haller (1986) found that students from rural schools achieved as well as students from urban schools. Kleinfeld et al. (1985) also did not find that high school location determined the quality of students' achievement on standardized tests. On average, literature showed that rural students perform about as well as urban students on national standardized tests.

2.3.2 Methods for Identifying Test Bias

In the literature, research has focused generally on two concepts for identifying test bias: differential item functioning (internal methods) and differential validity and differential prediction (external methods). Differential item functioning (DIF) is used to examine test bias at the item level. Item bias is one of the potential threats to the validity of any test that occurs when a test item is biased favoring one group of examinees over another. Differential prediction is used to examine whether the prediction equations obtained for the different groups of examinees are significantly different on the test scores (Linn, 1982; Shepard, 1987; Young, 2001).

In DIF analysis, individuals of the different subgroups are matched on a measure of ability, typically total test scores. The probabilities of a correct response on a particular item for individuals of the subgroups are then compared (Linn, 1993). The assumption of DIF is that examinees with comparable abilities should perform equally on each item in the test regardless of their group membership. DIF is said to occur when a test item is found to be substantially harder for one group than for another group, after having controlled for overall differences in the ability levels of the groups. DIF was not implemented in this study since the focus was not on the specific items of the college entrance exams.

In differential prediction analysis, evidence of bias or lack of bias is generally examined using the relationships between test and criterion scores for the different groups. When evidence is found by comparing the patterns of association between test scores and criterion variables for the respective groups, differential prediction or prediction bias is said to occur (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). Predictive bias refers to the systematic error in predicting the criterion variable for particular groups of students (House & Keeley, 1993). Differential prediction across gender and high school location groups was examined in this study.

2.3.3 Differential Validity or Differential Prediction

Differential validity and differential prediction are two important concepts that are considered by researchers examining prediction bias among test-takers. These two concepts appear to be very relevant to the issue of bias in selection decisions (Shepard, 1987). Linn (1982) described differential validity as the differences in the magnitude of the correlation coefficients or validity

coefficients for different groups of examinees. Young (2001) also stated that differential validity refers to the situation where a test is predictive for all groups but to different degrees. Differential validity is used to investigate whether the correlation between the predictor variables and the criterion variables are different for the different identifiable groups of examinees.

Differential prediction, on the other hand, occurs (according to the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1995) when "different algorithms are derived for the different groups and if the prediction leads to decisions regarding people from the individual groups that are systematically different from those decisions obtained from the algorithm based on the pooled groups" (p. 12). Linn (1978) defined differential prediction as the differences in the best-fitting regression lines or in the standard errors of estimate between groups of test-takers. Differences in regression lines are measured as differences in the slopes and/or intercepts. Linn (1982) claimed that differential prediction describes the fact that differences in prediction have more relevance when considering fairness in selection than do differential validity coefficients. Nevertheless, literature shows that both differential validity and differential prediction consistently demonstrate evidence of differences in the prediction equations across different groups of test-takers. Differential validity and differential prediction are obviously related but are not identical issues. Of the two issues, differential prediction is the more crucial because differences in prediction are more related to the issue of fairness in selection than differences in correlation (Linn, 1982a, 1982b).

In a psychometric sense, differential validity is important because it has relevance for the issues of test bias and fairness. Bias is described by Shepard (1982) as "invalidity, something that distorts the meaning of test results for some groups" (p. 26). Fairness refers to the

psychometric properties of the test and how the scores are used. Thus, if a test is differentially valid for different groups of examinees, it is consistently unfair to particular groups of examinees.

2.3.4 Research on Differential Validity and Differential Prediction

A number of studies have been conducted on differential validity or differential prediction, as well as on the combination of both on college admissions tests such as the SAT and GRE. Many researchers have reviewed differential validity and/or differential prediction studies (Breland, 1979; Burton & Ramist, 2001; Cleary, Humphreys, Kendrick, & Wesman, 1975; Duran, 1986; Linn, 1973; Morgan, 1989; Stanley & Porter, 1967; Wilson, 1983; Young, 2001). Breland (1979) conducted an extensive review of a number of studies on differential validity and differential prediction from 1964 to 1974. Predictors used in these studies included SAT mathematics scores, SAT verbal scores, and high school grades; the criterion variable used was first-year GPA. Breland suggested that, in terms of differential validity, the median values of the predictors for females were generally equal to or higher than that for males. For differential prediction, Breland found that there was overprediction in college performance for minority students.

Overall, literature on differential validity and differential prediction has also shown that admission test scores overpredict future performance for minority students (Cleary, 1968; Ramist, 1984; Young, 1991, 1994). Thomas (1972), in a study of differential prediction based on gender at 10 colleges, found that females' GPAs were underpredicted when a similar prediction equation was used to predict first-year GPA. Young (2001) also reviewed about 50 studies that examined differential prediction and/or differential validity in predicting future performance

across gender and ethnicity. He found consistent results across the various studies. His findings indicated that females' college performance was often underpredicted, and minority students' performance was overpredicted.

An explanation for this finding has been offered by many researchers. For example, Burton and Ramist (2001) explained that test scores do not necessarily underpredict females' future academic performance, but rather, the actual grades obtained by females are higher than predicted because females tend to enroll in easier courses. Some studies that have adjusted prediction equations for differences in college grading patterns have shown that the appearance of bias is indeed reduced or completely eliminated (Elliot & Strenta, 1988).

It is noteworthy to mention that differences across institutions, majors, and courses may alter the findings relative to differential validity and differential prediction in postsecondary institutions. Young (1993) comprehensively reviewed methods developed to adjust for grading differences. He concluded that in most differential validity and differential prediction studies the results are spuriously confounded if these factors are not accounted for. This should be taken into consideration when researchers interpret the results of differential validity or differential prediction. Elliott and Strenta (1988) showed that many studies have different results based on whether adjustments have been implemented to majors. Given the many conflicting results of differential validity and differential prediction studies, these issues continue to be of interest to many researchers, and they need to be investigated whenever it is plausible (Linn, 1994).

2.4 PERSISTENCE/RETENTION RESEARCH

A major challenge facing postsecondary institutions is the retention of their students. Seidman (2005) found that one-third of incoming freshmen in higher education leave the educational institution before completing their degree program. There have been a number of studies on predicting students' future academic success based on their high school GPA (Ishitani, 2005, 2006; Noble, Sawyer, 1997; Rosenbaum, 2004; Terenzini et al, 1996; Tinto, 1975, 2004; Van T. Bui, 2002; Waugh & Micceri, 1994), standardized test scores (Noble & Sawyer, 1997; Riehl, 1994; Rosenbaum, 2004; Waugh & Micceri, 1994), and difficulty of high school curriculum (Adelman, 1999, 2006; Ishitani, 2005, 2006; Pascarella, Terenzini, 1980; Tinto, 1975; Warburton et al., 2001), but few studies have been conducted in using these factors to predict students' retention and persistence to graduation.

Cognitive factors such as high school GPA and standardized test scores have been determined to have some level of validity in predicting students' persistence to graduation. Rosenbaum (2004) stated, "(T)he easiest-to-use predictor of a student's likelihood of graduating from a college is still their high school grade point average" (p. 2). Ishitani (2005) also concluded that "High school class rank and high school academic intensity had significant effects on college attrition behavior" (p. 18). Because high school class rank is based on student GPAs, this shows the importance of a student's high school GPA on his/her persistence to graduation from college. Moreover, Peterson, Casillas, and Robbins (2006) found for future academic success higher emphasis on is placed on the ACT or SAT and high school GPA for identifying students who are at risk of dropping out of higher education. In a more recent study, Friedman and Mandel (2010) reported that students' SAT scores and high school GPA are significantly

related to both cumulative GPA and retention after the first year. Hezlett et al. (2001), in a meta-analysis study on the SAT, also concluded that "Individuals with higher SAT scores are more likely to have good study habits, remain in college, and complete their degrees" (p. 13). In the same stream, Waugh and Micceri (1994) stated, "Students with higher high school GPAs have much higher retention/graduation rates" (p. 4). They contended that students with a high school GPA have a retention/graduation rate at least nine percent higher than students with less than 2.4 high school GPA.

Although cognitive factors have achieved wide acceptance as tools to predict future academic success, some researchers have challenged this notion and criticized them as insufficient indicators to address students' persistence. Ting (1998), for example, stated that ACT scores were not significant predictors of persistence to graduation. Lohfink and Paulsen (2005) and Munro (1981) also found that college entrance test scores do not have a direct tie to the retention of a student in higher education institutions.

Researchers proposed different models to study students' retention in postsecondary institutions (Alfred, 1973; Astin, 1977; Pascarella & Terenzini, 1980; Tinto, 1975). However, the Tinto (1975) model appears to be the most widely accepted to explain college students' retention and persistence. The model focuses on the adaptation of students to the college environment. In his model, he explained that retention depends on students' background, goal and institutional commitment, and degree of academic as well as social integration. Littleton (2001) examined students who have persisted at small colleges. His results indicated that persisting students listed faculty influence, campus involvement, support from family, peer relationships, and a positive attitude as important contributors to their success and retention. Noticeably, many faculty in Yemen do not encourage nor welcome a friendly relationship with their students for many

reasons. Some educators explained that those faculty are not confident of their ability to answer all students' questions and do not want to be embarrassed by their inability to do so (Sa'ad, 1998; Alkhateeb, 1992). Another reason is that some faculty think that students should ask questions only in the classroom. In addition, due to the fact that in Yemen high school students come to colleges and universities from completely separated schools; male students come from schools that are operated by male personnel and female students come from schools that are operated by female personnel. Their participation and interaction, especially females, has been limited in the existence of the opposite sex (UNESCO, March 2006).

Tinto (1982) discussed how the students' level of academic and social integration impacts their reasons for retention. He stated that institutions must take into account the longitudinal nature of student persistence. Students enter higher education with a variety of educational interests, values, skills, goals, and commitments. The influences on students' departure decisions vary throughout their college careers. The first college semester has been considered the most important time period for colleges and universities to impact student persistence, but students can choose to leave an institution at any time during their college career.

Adelman (1999) identified that the bottom line that matters to all parties is students' persistence to graduation. He developed a longitudinal study that focused on high school GPA, college transcripts, and surveys of the students. The results he found established that there was a correlation between bachelor degree attainment and curriculum, test scores, and high school GPA. However, predicting retention using high school GPA and standardized test scores is still questionable (Armstrong, 1999; King, Rasool, Judge, 1994). From the literature it does not appear that either side has produced conclusive evidence to support or disprove the validity of

high school GPA and college entrance test (e.g., ACT/SAT) scores as valid predictors of persistence to graduation.

3.0 METHODOLOGY

This chapter provides a detailed description of the methodology used in this study. It presents the purpose of the study, research design, sample, data collection procedures, variables, and data analysis techniques utilized in the study.

3.1 PURPOSE OF THE STUDY

At the time of the establishment of the only two universities in Yemen (Sana'a University and Aden University) in 1970, high school GPA was the only requirement for admission at these universities. There was a seat for every student who wanted to continue postsecondary education. The idea at that time was to educate as many Yemeni citizens as possible to replace the foreign teachers. After the unification of North and South Yemen in 1990, high school GPA was used, for a short period of time, as the only criterion for admission at all educational institutions beyond high school. However, when government officials noticed that the number of students exceeded the capacity of both universities, they decided to raise the high school GPA percentage score required to be admitted to university. The required percentage score varied depending on each college. Colleges of Science, Medicine, and Engineering and some other departments

required entrance test scores in addition to high school GPA to determine the admission of a student.

The purpose of this study was to investigate the predictive validity of both high school grade point average and college entrance test scores used as criteria in the admission process to postsecondary institutions in the Republic of Yemen. The study intended to examine if adding college entrance test scores to high school GPA would increase the prediction power as measured by first-year college grade point average and four-year cumulative grade point average. In addition, the differential predictive validity of high school GPA and college entrance test scores was examined across gender and high school location. The relationship between students' persistence in the four years of college and the predictor variables, high school GPA and college entrance test scores, was studied as well. The study also examined the internal consistency reliability of college entrance test scores. The college entrance tests consisted of different subject areas. It was necessary to examine if the subtests as well as the overall tests were truly reliable and the extent to which any errors in measurement may influence the analyses.

3.2 RESEARCH DESIGN

This research was designed to examine the predictive validity of high school grade point average and college entrance test scores in predicting students' college academic success, as measured by first-year college grade point average and four-year cumulative grade point average as well as students' persistence, as measured by a dichotomous variable (graduated/not graduated). High school GPA and college entrance test scores were chosen as predictors (independent variables) and first-year college GPA, four-year cumulative GPA, and students' persistence were selected to serve as criteria (dependent variables) in this study. In order to answer the main research questions, multiple and logistic regression analyses were employed. Internal consistency analysis was employed to examine the reliability coefficient of college entrance test scores.

3.3 RESEARCH SETTING

This study was conducted with 2010 graduates at two universities in Yemen, Hodeidah University and Ibb University. Hodeidah University was established in the Hodeidah governorate as an official university in 1996. Before its foundation, its associated College of Education had already been established in 1988 and, at that time, was affiliated with Sana'a University. Hodeidah University has 11 colleges currently. Ibb University was also established in 1996 in the governorate of Ibb and has eight colleges (Appendix E). Students enrolled in these universities come from different parts of the country. Both universities provide education to males and females in the same buildings and classrooms.

3.4 POPULATION AND SAMPLING

The sample included 881 cohort students who entered the two universities (Hodeidah and Ibb) in 2006/2007. Out of this pooled sample size, 764 students graduated after the norm-time of four years in college. The information and records of all students who were admitted in 2007 were obtained from the admission files of the universities' registrar offices. The study was conducted using information obtained exclusively from the institutions' databases. The data, which included high school grade point average, college entrance test scores, major in colleges, first-year college grade point average for those who finished their first year of college, and four-year cumulative grade point average for those who made it to their fourth year of college, was directly extracted from the files of students in these universities. Data for persistence was coded as '1' for a student who graduated from college in 2010 and '0' for a student who did not graduate. Permission to collect data was obtained from the presidents of the two universities.

The sample comprised of 52.33% females and 47.67% males (Table 1). In the study 36.44% of the students came from rural high schools and 63.56% came from urban high schools (Table 2). Frequencies of students for each college by gender and high school location are presented in Appendix (H).

Table 1. Frequency Distribution of Participants by Gender

Gender	Frequency	Percentage
Female	461	52.33
Male	420	47.67
Total	881	100.00

Table 2. Frequency Distribution of Participants by High School Location

Location	Frequency	Percentage
Urban	560	63.56
Rural	321	36.44
Total	881	100.00

3.5 STUDY VARIABLES

3.5.1 Predictor Variables

The main independent variables utilized in this study are high school grade point average and college entrance test scores. Each of these variables is described below. Other variables include gender, high school location, and college major.

3.5.1.1 High School Grade Point Average

High school grade point average refers to the mean percentage score over all subjects taken in the twelfth grade. It is obtained from the general secondary examinations that are constructed by the Ministry of Education and administered to all twelfth graders across the country at the end of the school year. A student should get a minimum percentage score of 50% in each subject to pass the general secondary examinations and a minimum overall percentage score of 70% to be able to enroll in higher education institutions (MoE, 2010).

3.5.1.2 College Entrance Tests

The college entrance tests are exams designed to measure a student's readiness for future academic success. The college entrance tests are composed of a combination of different subtests. These subtests are English, Physics, Chemistry, Biology, and Mathematics. These tests are prepared by expert professors at the specific departments in the Colleges of Education. For example, the English department prepares the English test and the physics department prepares the physics test. Each subtest consists of a range of 20–30 multiple-choice items of three or four

alternatives. The Colleges of Arts and Colleges of Education use only the English subtest for admission. These tests are made of 30 multiple-choice items of three alternatives, distributed over three main parts: reading comprehension, grammar, and vocabulary. Responses to the items are scored dichotomously as either 1 (*correct*) or 0 (*incorrect*), with no penalty for incorrect responses. Each college has a different combination of these subject areas (Appendix F). The college entrance tests are administered around the same time (between August-September) every year. The scores of the college entrance tests are used along with high school grade point average to make significant decisions about students' potential to succeed in their college studies. The maximum percentage score on the entrance test is 100%. Despite the fact that the pass-point is 50%+, there is no restriction to this limit, particularly when the required number may urge the admission committee to go below this percentage.

Admission to most colleges is based on both high school grade point average and college entrance test scores. The weight is 50% for each of them. For example if a student gets 80% in high school GPA and 70% in college entrance test scores, then his/her composite score is 75%. Students' scores are arranged high to low, and the required number of students in college is selected by the admission officers. For college entrance tests that have different subtests, the same method is applied. For example, a test has three subtests (English, Math, and Physics) at the College of Computer Science and Engineering. Each test is conducted on a different day. If a student scores 70% in English, 60% in Math, and 75% in Physics, then the student's total percentage score on the three tests is averaged and compared to a perfect score of 100%, meaning that the student's total percentage score is 68.33% on this given exam.

3.5.2 Criteria Variables

Robbins and colleagues (2004) identified that students' college grade point average and persistence are two major domains of college student outcomes. Because first-year college grade point average provides only a short-term indicator of college performance, the present study included four-year cumulative GPA and students' persistence for the same sample in order to examine the relative contribution of high school GPA and college entrance test scores in predicting longer-term college performance. The outcome variables used in this study are first-year college GPA, four-year cumulative GPA, and students' persistence. The college grade point average refers to the average grade point that students obtain based on a scale with a maximum of 100 points and a minimum of 0 points (Appendix G). It is calculated for each student every semester. A description and rationale for each of the criteria variables is presented below.

3.5.2.1 First-year College Grade Point Average

Research shows that first-year college GPA is the most commonly used criterion variable in the predictive validity studies of college admission procedures (Wilson, 1983). Pascarella and Terenzini (1991) stated that "First-year grades are probably the single most revealing indicator of . . . successful adjustment to the intellectual demands of a particular college's course of study" (p. 388). Camara and Echternacht (2000) found that first-year college GPA is the most frequently used criterion in predictive studies. They noted that first-year college GPA is favored because it is a well-defined criterion. First-year college GPA scores are easily retrieved from university records, and it is available relatively soon after students finish high school.

Willingham (1985) found that high school grades are highly correlated with first-year GPA as well as cumulative GPA over time. Thus, the first-year college GPA could be deemed as a good predictor of subsequent years.

3.5.2.2 Four-year Cumulative Grade Point Average

Although many researchers are concerned with the prediction of academic success over time, there are few studies using four-year cumulative grade point average as a criterion variable of students' academic performance. Young (1989) claimed that

> The four-year cumulative GPA has greater face validity and greater content validity as a measure of overall achievement than first-year grades alone. In addition, the cumulative GPA should exhibit better reliability than the first-year GPA since it represents a larger sample of each student's behavior.

Camara and Echternacht (2000) also stated that "The rationale for considering cumulative GPA as an indicator of success in college is that it encompasses the entire scholastic performance of a student at a college." Four-year cumulative GPA was used in this study because it appeared to provide a total achievement of student academic performance.

3.5.2.3 Student Persistence

The ultimate goal for students going to college is persistence to graduation (Adelman, 1999). Although research findings suggest that persistence in higher education is heavily influenced by nonacademic factors such as finances, family and social considerations, and campus involvement, (Pascarella & Terenzini, 1980; Astin, 1977; Tinto, 1975), high school GPA and standardized test scores were concluded to have some validity in predicting students' persistence

(Ishitani, 2005; Pike and Saupe, 2002; Peterson, Casillas, and Robbins, 2006; Friedman and Mandel, 2010). Pike and Saupe (2002) noted that high school grades are highly correlated with students' persistence and graduation. Friedman and Mandel (2010) also suggested that high school GPA and standardized test scores (e.g., SAT) are significantly related to both cumulative GPA and student retention. Student persistence is operationalized as whether or not a student completes a degree program within a specified time (Astin, 1977, p. 7). In this study dichotomous data was used for the persistence variable. If a student graduated, he/she was coded "1" and if not, was coded "0".

3.6 DATA ANALYSIS

The data analysis for this study included both descriptive and inferential statistics. Descriptive statistics were computed for predictor variables (high school GPA and college entrance test scores) and for the criterion variables (first-year college grade point average, four-year cumulative grade point average, and students' persistence). Frequencies for demographic variables (gender and high school location) were also tabulated. All analyses were performed separately. This is due to the fact that college entrance tests are college-dependent. Data analyses were therefore carried out separately for the nine different colleges at the two universities.

Internal consistency reliability estimates for college entrance test scores were examined in this study using Cronbach's coefficient alpha measure that was proposed by Cronbach (1951). For composite entrance test scores, the stratified alpha coefficient that was introduced by Cronbach, Schoneman, and McKie (1965) was utilized. Stratified alpha provides a better

estimate than coefficient alpha in the case of multiple subtests (Osburn, 2000). Hierarchical multiple regression analyses were used to answer the research questions. These analyses evaluated whether high school GPA was an accurate predictor in predicting college academic success and whether adding a college entrance test scores improved the prediction validity as measured by first-year college GPA and four-year cumulative GPA. High school GPA and college entrance test scores were further examined for differential prediction across subgroups (gender and high school location). Logistic regression analysis was employed in the analysis of four-year persistence to graduation with persistence defined as a dichotomous outcome variable (graduated/not graduated). The hypotheses in the current study were tested at the .05 level of significance. All the analyses were conducted using Statistical Analysis System (SAS, version 9.2) software.

3.6.1 Predictive Validity Analyses

First-year college GPA and four-year cumulative GPA were regressed separately on the same set of predictor variables. High school GPA and college entrance test scores were entered into the hierarchical regression models in that order. This sequence was based on a theoretical framework derived from previous predictive validity research. Hierarchical multiple regression analysis provides the proportion of explained variance, also known as the *coefficient of determination* or R^2, in the criterion variable that is accounted for or explained by the predictor variables (high school GPA and college entrance test scores) in each model in the regression and the change in R^2 in the following model. Adjusted R^2 value is a value that takes into account the number of variables in the model and the number of observations or participants the model is based on. One

can test the significance of difference of two R^2's by the *F*-test to determine if adding an independent variable to the model helps significantly. Using hierarchical regression, one can see how most variance in the dependent variable can be explained by one or a set of independent variables, over and above that explained by an earlier variable(s). The results also provide unstandardized and standardized regression coefficients in each model of the regression analyses. Standardized regression coefficients, also known as "beta weights," are useful in situations where the predictor variables involve widely different measurement scales. Regression coefficients show the amount that the dependent variable increases (e.g., first-year college GPA) when the independent variable increases one unit, all other factors held constant. The estimates (b coefficients and constant) can be used to construct a prediction equation. In regression analysis, the prediction equation is often presented as follows:

$$Y = a + b_1X_1 + b_2X_2 + e$$

where 'Y' is the predicted or dependent variable, 'X_1' is the score on the first predictor variable (high school GPA), and 'X_2' is the score on the second predictor variable (college entrance test scores). 'a' is the constant or intercept, representing the amount that the dependent variable will be when all the independent variables are zero. The regression coefficients (b_1, b_2) represent the amount the dependent variable changes when the corresponding independent variable changes one unit, all other factors held constant. Finally, '*e*' stands for prediction error or residual, the difference between observed value and predicted value of the dependent variable. The main aim of regression procedures is to minimize the amount of prediction errors so as to maximize the linear relationship between the independent variable and the dependent variable (Kerlinger & Pedhazur, 1973). The use of multiple regression helps explain the variance of a dependent variable by estimating the contributions of one or two independent variables to this variance.

Prior to conducting hierarchical multiple linear regression analyses, an intercorrelation matrix using Pearson product moment correlations was examined to determine the extent to which each of the independent variables was related to the dependent variable. The assumptions underlying regression were also examined to assess the appropriateness of the regression models (Mertler & Vannatta, 2002). Normality, linearity, and outliers were screened using histograms, scatterplots, and Q-Q plots. Normality and linearity were also assessed using Shapiro-Wilk test and Box-Cox transformation, respectively. The assumption of homoscedasticity was examined as well. The scatterplot is an effective way to check for homoscedasticity (that is the error terms or residuals along the regression line are equal). Homoscedasticity was also tested using White and Breusch-Pagan tests.

3.6.2 Differential Prediction Analyses

Testing for differential prediction involves a test of equality of standard errors of estimate, followed by a test of equality of regression slopes, and then a test of equality of regression intercepts across identifiable subgroups (Gulliksen & Wilks, 1950). The standard errors of estimate measures the amount of errors involved in the prediction of the criterion variable from the predictor variable (Jensen, 1980). The differences in regression slopes mean that the test predicts performance better for one group than for another, while the differences in regression intercepts mean that the members of one group tend to obtain lower predicted scores than the members of another group who are of equal ability (Sackett & Wilk, 1994).

3.6.3 Logistic Regression Analyses

For the logistics regression analyses, the procedure allows for independent variables to predict a dichotomous outcome. In this study one of the criterion variables (persistence) is dichotomous (whether a student graduated or not within the norm-time of four years in college) and the predictor variables are high school GPA and college entrance test scores. Logistics regression analysis uses the maximum likelihood estimation rather than the least-squares estimation to obtain the estimates of the model parameters. The logistic regression models the relationship between independent variables and criterion variable through the logit function. The following equation represents the logit equation for the persistence model in the current investigation:

$$Log\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + e$$

where $Log\left(\frac{p}{1-p}\right)$ equals the natural log of the odds ratio, and the odds ratio is the probability of persistence divided by the probability of dropping out:

$$Odds\ Ratio\ for\ Persistence = \frac{P(Persistence)}{P(Dropout)}$$

If the variable is significant, then the logit coefficient suggests a significant increase or decrease in the odds ratio. The probability that a student persists to graduation equals the following equation:

$$P(Persistence) = \frac{Odds\ Ratio}{(1\ +\ Odds\ Ratio)}.$$

The effectiveness of logistic regression models was examined by using the likelihood ratio test. This test examines if the logistic model is more effective than the null model, the intercept-only model. The statistical significance of individual regression coefficients (i.e., *B*s) was tested using the Wald chi-square statistic. In addition, the goodness-of-fit test was examined using the Hosmer–Lemeshow (H–L) statistic which should not be significant to imply that the logistic model fits the data well. Another statistic of goodness-of-fit that was examined is Nagelkerke R^2. The intuition of Nagelkerke R^2 is to calculate the improvements of the prediction by the model from the null model. The log-likelihood is used for the calculation of Nagelkerke R^2. The interpretation of Nagelkerke R^2 is not the same as in the ordinary least squares (OLS) regression, but they can be interpreted as an approximate variance in the outcome accounted for by the predictor variables.

4.0 RESULTS

The study was intended to examine the predictive validity of high school grade point average and college entrance test scores, the only two criteria used in the admission process to postsecondary institutions in Yemen. The study also examined if the addition of college entrance test scores to high school GPA would increase the prediction power as measured by first-year college grade point average and four-year cumulative grade point average for each of the participating colleges. The differential predictive validity of high school GPA and college entrance test scores was examined across gender and high school location. The relationship between students' persistence in the four years of college and the predictor variables, high school GPA and college entrance test scores, was studied as well. Finally, the study also examined the internal consistency reliability of college entrance tests scores.

This chapter presents the results of the data analyses. It is presented according to the order of the research questions. The results are organized into six sections. The first section presents descriptive statistics for the sample and the variables in the study. The second section provides the reliability coefficient estimates of the college entrance test scores obtained from the internal consistency analyses using Cronbach's alpha for the individual tests and stratified alpha procedures for the composite tests. The third section presents the findings from the correlation and regression analyses of the predictor variables and first-year college GPA. Differential

prediction of the predictor variables across gender and high school location is reported in the fourth section. The fifth section focuses on the predictive validity of high school GPA and college entrance test scores over time as measured by the four-year cumulative grade point average. The final section presents the relationship between the predictor variables and students' persistence in college as measured by whether a student graduated in the norm-time of four years of study in college. Implications of the findings are discussed in Chapter 5.

4.1 DESCRIPTIVE STATISTICS

Descriptive statistics (frequency, mean, and standard deviation) were computed first for the entire sample in the study (Table 3) and then by gender and high school location (Tables 4 and 5). The overall mean percentage score of high school grade point average across all colleges was 82.19 (*SD* = 6.40). The overall mean percentage score of college entrance tests was 78.32 (*SD* = 10.76). This means that on average students received 82.19% and 78.32% of points in high school GPA and college entrance test scores, respectively.

When looking at the means for the same variables by gender and high school location, similar results were observed for the different subpopulations. For female students, the mean percentage score of high school GPA was 82.24 (*SD* = 6.37), while it was 82.13 (*SD* = 6.44) for male students. The mean percentage score of college entrance test for female students was 79.73 (*SD* = 9.98) and for male students was 76.78 (*SD* = 11.37). For students coming from rural high schools, the mean percentage score of high school GPA was 81.65 (*SD* = 6.40), while it was 82.50 (*SD* = 6.38) for students coming from urban high schools. The mean percentage score of

college entrance test for students coming from rural high schools was 78.63 (*SD* = 10.42), and for those coming from urban high schools, the mean percentage score was 78.14 (*SD* = 10.94).

For the criterion variables, the students' mean score on first-year college grade point average (*M* = 74.88, *SD* = 7.02) was very close to that of four-year cumulative grade point average (*M* = 74.48, *SD* = 7.56). Based on gender, female students achieved a higher mean score on both first-year college GPA (*M* = 75.61, *SD* = 7.05) and four-year cumulative GPA (*M* = 75.00, *SD* = 7.89) than male students (*M* = 74.08, *SD* = 6.90) and (*M* = 73.93, *SD* = 7.17), respectively. When looking at the descriptive statistics by high school location, urban students received a higher mean score on first-year college GPA (*M* = 75.03, *SD* = 7.04) and four-year cumulative GPA (*M* = 74.51, *SD* = 7.53) than rural students (*M* = 74.62, *SD* = 6.99) and (*M* = 74.43, *SD* = 7.62), respectively.

Table 3. Descriptive Statistics of the Variables in the Study across all Colleges

Variable	N	Mean	SD	Minimum	Maximum
High School GPA	881	82.19	6.40	68.34	96.87
College Entrance Test Scores	881	77.56	11.14	42.41	99.38
First-year GPA	881	74.88	7.02	53.64	91.76
Four-year Cumulative GPA	881	74.48	7.56	53.67	92.36
Graduated (1 = Yes, 0 = No)	764	0.87*	0.34	0.00	1.00

* This value corresponds to the proportion of graduates.

Table 4. Descriptive Statistics of the Variables in the Study by Gender across all Colleges

Variable	Gender	N	Mean	SD	Minimum	Maximum
High School GPA	F	461	82.24	6.37	68.34	96.87
	M	420	82.13	6.44	68.76	96.78
College Entrance Test Scores	F	461	78.78	10.62	46.00	99.38
	M	420	76.22	11.55	42.41	99.00
First-Year GPA	F	461	75.61	7.05	56.24	91.76
	M	420	74.08	6.90	53.64	89.63
Four-year Cumulative GPA	F	461	75.00	7.88	55.39	91.08
	M	420	73.93	7.17	53.67	92.36
Graduated (1 = Yes, 0 = No)	F	396	0.86	0.35	0.00	1.00
	M	368	0.88	0.33	0.00	1.00

Table 5. Descriptive Statistics of the Variables in the Study by High School Location across all Colleges

Variable	Location	N	Mean	SD	Minimum	Maximum
High School GPA	R	321	81.65	6.40	68.34	96.87
	U	560	82.50	6.38	68.76	96.78
College Entrance Test Scores	R	321	77.68	11.09	47.38	99.00
	U	560	77.50	11.18	42.41	99.38
First-Year GPA	R	321	74.62	6.99	55.46	91.76
	U	560	75.03	7.04	53.64	90.60
Four-year Cumulative GPA	R	321	74.43	7.62	54.42	90.74
	U	560	74.51	7.53	53.67	92.36
Graduated (1 = Yes, 0 = No)	R	266	0.83	0.38	0.00	1.00
	U	498	0.89	0.31	0.00	1.00

Tables 6 and 7 show the means and standard deviations of high school grade point average and college entrance test scores for each college in the study by gender and high school location. The means percentage score of high school GPA varied based on students' majors. The highest mean percentage score was at the College of Medical Science at Hodeidah University (M = 88.14, SD = 4.29), while the lowest mean percentage score was at the College of Arts at Hodeidah University (M = 75.30, SD = 3.86). The mean percentage score of high school GPA for hard-science colleges (College of Computer, College of Dentistry, College of Engineering

and Architecture, and College of Medical Science) were high due to the fact that these colleges require a high school GPA of 80% and above. The results also showed that the mean percentage of college entrance test scores for students at the Colleges of Arts and Education at both universities were above 80%. This may be because these colleges use only the English subtest (not a composite test) for the entrance exam to admit students into the English departments. The standard deviations of College entrance test scores for some colleges (College of Arts, College of Dentistry, College of Engineering and Architecture, and College of Sciences at Ibb University) were slightly higher than the others due to the fact that these colleges had to go below 50% of College entrance test score to select their prospective students. Admitting students who scored below 50% resulted in a larger variance in the scores.

Based on gender, the mean percentage score for most colleges showed that female students scored higher than male students on both predictors, high school grade point average and college entrance test scores. This is similar to the overall results across all colleges. Similarly, the mean percentage score for students who studied in urban high schools was higher than the mean percentage score for students who studied in rural high schools on both high school GPA and college entrance test scores at most colleges.

Table 6. Descriptive Statistics of the Predictor Variables for each College by Gender

College	Gender	N	High School GPA		College Entrance Test Scores	
			M	SD	M	SD
Arts (Hodeidah)	F	30	76.73	5.39	80.38	7.44
	M	15	73.87	2.33	80.59	5.65
Arts (Ibb)	F	80	77.00	4.48	84.19	12.08
	M	86	76.69	3.96	80.38	13.93
Computer Science and Engineering (Hodeidah)	F	31	86.99	4.74	81.90	6.97
	M	65	86.44	4.14	81.57	7.98
Dentistry (Hodeidah)	F	50	86.27	3.22	75.82	6.66
	M	17	85.69	3.22	70.91	8.75
Dentistry (Ibb)	F	42	86.96	1.51	74.82	11.43
	M	26	86.84	1.14	68.37	13.16
Education (Hodeidah)	F	79	78.88	4.77	87.86	5.24
	M	28	78.64	5.61	88.22	4.34
Engineering and Architecture (Ibb)	F	20	81.86	1.34	77.69	10.06
	M	31	82.36	1.80	74.83	10.13
Medical Science (Ibb)	F	91	88.32	4.10	75.99	6.40
	M	96	87.96	4.48	75.36	6.29
Sciences (Ibb)	F	38	75.81	2.63	71.76	9.58
	M	56	76.07	3.23	68.15	11.18

Table 7. Descriptive Statistics of the Predictor Variables for each College by High School Location

College	HS Location	N	High School GPA		College Entrance Test Scores	
			M	SD	M	SD
Arts (Hodeidah)	R	21	75.19	3.81	78.62	4.16
	U	24	76.29	5.50	82.05	8.28
Arts (Ibb)	R	66	76.98	4.16	81.69	13.48
	U	100	76.75	4.26	82.57	13.02
Computer Science and Engineering (Hodeidah)	R	38	87.08	4.35	81.37	7.46
	U	58	86.32	4.31	81.88	7.80
Dentistry (Hodeidah)	R	11	85.82	3.89	72.54	8.41
	U	56	86.19	3.09	74.97	7.32
Dentistry (Ibb)	R	25	86.61	1.17	72.83	12.51
	U	43	87.09	1.46	72.08	12.52
Education (Hodeidah)	R	44	77.99	4.78	86.49	4.54
	U	63	79.40	5.07	88.98	5.09
Engineering and Architecture (Ibb)	R	23	82.27	1.53	79.55	9.38
	U	28	82.08	1.74	73.00	9.85
Medical Science (Ibb)	R	63	87.78	4.65	75.00	6.71
	U	124	88.31	4.10	76.01	6.13
Sciences (Ibb)	R	30	75.91	2.80	70.92	10.18
	U	64	75.99	3.10	68.99	10.90

4.2 RELIABILITY OF COLLEGE ENTRANCE TEST SCORES

The reliability of the college entrance test scores was investigated using the internal consistency measure. Internal consistency estimates the magnitude of errors associated with content sampling and was used as a method for estimating reliability because estimates of internal consistency can be obtained from tests (like college entrance tests) conducted at one point in time. There are five college entrance tests at each of the universities (Hodeidah and Ibb). These tests are English, Physics, Chemistry, Biology, and Mathematics. Each test is prepared by experts in the specific department at the Colleges of Education. Each college then selects the test(s) that they want to use with their prospective students.

Reliability coefficients of the college entrance test scores were examined individually and in composite form within each college in the two universities (see Tables 8 and 9). The reliability coefficients of the individual college entrance test scores were estimated using Cronbach's alpha measure. The reliability of the composite tests within each college was estimated using the stratified alpha coefficient measure. This measure was introduced by Cronbach, Schoneman, & McKie (1965) and is used to compensate for the different dimensions of the subtests.

The reliability coefficients of the composite college entrance test scores within each college ranged from 0.85 to 0.88 at Hodeidah University. The highest reliability coefficient was for the composite test (English, Physics, Math) at the College of Computer Science and Engineering, while the lowest was for the composite test (English, Chemistry, Biology) at the College of Medical Science. The reliability coefficients of the composite college entrance test scores within each college at Ibb University ranged from 0.81 to 0.89. The highest reliability

coefficient was for the composite test (English, Physics, Chemistry, Biology) at the College of Dentistry, while the lowest was for the composite test (English, Physics, Math) at the College of Engineering and Architecture. The subtests had acceptable reliability coefficients when examined separately at both universities. The reliability coefficients ranged from 0.77 to 0.94 at Hodeidah University and from 0.79 to 0.92 at Ibb University. In general, all of the tests, individually and in composite, appeared to be reliable.

Table 8. Reliability Coefficients for College Entrance Test Scores within each College at Hodeidah University

College	# of Items	M	SD	Stratified α	Subtest	# of Items	M	SD	α
Arts	NA	NA	NA	NA	English	30	27.47	4.65	.94
Computer Science and Engineering	75	63.11	8.03	.88	English	30	24.22	5.53	.88
					Physics	20	16.85	3.16	.78
					Math	25	22.04	3.49	.82
Dentistry	100	79.64	10.01	.87	English	30	23.63	5.43	.86
					Physics	20	17.06	3.54	.84
					Chemistry	25	20.12	3.74	.77
					Biology	25	18.84	5.31	.87
Education	NA	NA	NA	NA	English	30	19.77	7.21	.91
Medical Science	80	68.10	7.83	.85	English	30	25.80	4.85	.85
					Chemistry	25	20.56	4.07	.78
					Biology	25	21.74	3.97	.86

NA = Not Applicable

Table 9. Reliability Coefficients for College Entrance Test Scores within each College at Ibb University

College	# of Items	M	SD	Stratified α	Subtest	# of Items	M	SD	α
Arts	NA	NA	NA	NA	English	30	23.80	5.74	.89
Engineering and Architecture	75	53.20	7.77	.81	English	30	19.49	6.11	.87
					Physics	20	13.16	4.40	.82
					Math	25	20.55	3.86	.79
Dentistry	100	67.32	11.60	.89	English	30	19.87	7.19	.92
					Physics	20	15.63	4.08	.84
					Chemistry	25	15.81	6.07	.90
					Biology	25	16.01	5.57	.88
Sciences	80	61.60	8.94	.87	English	30	24.53	4.74	.83
					Chemistry	25	16.57	5.44	.87
					Biology	25	20.49	4.28	.83

NA = Not Applicable

4.3 PREDICTIVE VALIDITY AS MEASURED BY FIRST-YEAR COLLEGE GRADE POINT AVERAGE

In this study the predictive validity of high school grade point average and college entrance test scores as measured by first-year college grade point average was examined for each college separately. Hierarchical regression analysis was implemented to examine the predictive validly of the predictor variables. Prior to performing the regression analyses, the data were evaluated to find out whether they met the assumptions of regression. Normality of residuals was evaluated through the assessment of histograms and Shapiro-Wilk test. Residuals are the differences between the observed and predicted values by the regression equation. Residuals thus represent errors which should be normally distributed for each set of values of the independent variables. Q-Q plots were examined for outliers because normality tests are biased by an outlier. The data met the assumption of normality as evidenced by histograms that approximated normality and nonsignificant Shapiro-Wilk statistics.

The linearity of the relationship between the predictor variables, high school GPA and college entrance test scores, and the criterion variable, first-year college GPA was assessed through the examination of scatterplots. Residual plots were also examined for the linear relationship between the residual values and the predicted scores. Furthermore, an objective measure, Box-Cox transformation, was also used to test for linearity. The linearity assumption is important because regression analysis only tests for a linear relationship between the independent variables and the dependent variable. Homoscedasticity assumption was also examined using White and Breusch-Pagan tests. The results indicated that the data met all the regression

assumptions after the deletion of influential outliers. The number of cases that were deleted ranged from 3 to 10 cases across the colleges. The outliers were not included in the all subsequent analyses.

The correlation coefficients (r_{xy}) between high school grade point average and first-year college grade point average and college entrance test scores with first-year college grade point average are presented in Table 10. The results indicated that there was significant low-to-moderate correlations between high school GPA and first-year college GPA at all colleges except one, the College of Engineering and Architecture at Ibb University, $r_{xy} = 0.22$, $p = .13$. The results also showed that there were significant moderate correlations between college entrance test scores and first-year college GPA at all colleges except two, the College of Engineering and Architecture, $r_{xy} = 0.21$, $p = .13$ and College of Sciences, $r_{xy} = 0.06$, $p = .59$ at Ibb University. The strongest significant relationship between high school GPA and first-year college GPA was at the College of Arts at Ibb University, $r_{xy} = 0.66$, $p < .01$; while the weakest significant relationship was at the College of Medical Science at Hodeidah University, $r_{xy} = 0.24$, $p < .01$. However, the strongest significant relationship between college entrance test scores and first-year college GPA was at the College of Medical Science at Hodeidah University, $r_{xy} = 0.71$, $p < .01$; while the weakest significant relationship was at the College of Dentistry at Ibb University, $r_{xy} = 0.41$, $p < .01$. In general and across all colleges, the magnitude of the relationship between the predictor variables and first-year college GPA was higher for college entrance test scores than for high school GPA. However, the correlation coefficient of the relationship between college entrance test scores and first-year college GPA for the College of

Sciences at Ibb University was found to be 0.06. This is an unusual finding that needs to be further analyzed.

The correlation coefficients of the relationship between high school GPA and college entrance test scores with first-year college GPA were also estimated for each college by gender and high school location and are presented in Tables 11 and 12. The analyses by gender revealed conflicting results. The correlations between the predictor variables and first-year college GPA were found to be not significant for male students at many colleges. The same findings were observed for rural and urban students. This may be due to the small sample sizes and the restriction in range for some subgroups. Scores are considered restricted because admitted students tend to have a narrower range of scores than the larger applicant pool that includes the students who were not selected. When correlation coefficients are estimated for subgroups, range of scores may even be more narrower. Restriction of range of test scores affects the estimated magnitude of the relationship between test scores and the criterion variable. Givner and Hynes (1979) demonstrated that the more restricted the distribution of test scores, the more the relationship between the test scores and the criterion variable will be underestimated. Restriction of range was not corrected in this study because doing so requires that the unrestricted variance of predictor variables be known. Data for students who did not meet the cut-off for high school GPA required by the particular colleges were not available and data for students who were not selected into the colleges were discarded by those colleges. Thus the necessary data were not available to correct for restriction of range.

Table 10. Correlation Coefficients between Predictor Variables and First-year College GPA

No.	College	N	HSGPA (r_{xy})	# of Subtests	CETS (r_{xy})
1	Arts (Hodeidah)	45	.38*	1	.43**
2	Arts (Ibb)	166	.66**	1	.44**
3	Computer Science and Engineering (Hod.)	96	.27**	3	.53**
4	Dentistry (Hodeidah)	67	.33**	4	.46**
5	Dentistry (Ibb)	68	.35**	4	.41**
6	Education (Hodeidah)	107	.40**	1	.45**
7	Engineering and Architecture (Ibb)	51	.22	3	.21
8	Medical Science (Hodeidah)	187	.24**	3	.71**
9	Sciences (Ibb)	94	.51**	3	.06

* < .05, ** < .01. HSGPA = High School Grade Point Average, CETS = College Entrance Test Scores

Table 11. Correlation Coefficients between Predictor Variables and First-Year College GPA by Gender

No.	College	Gender	N	HSGPA (r_{xy})	CETS (r_{xy})
1	Arts (Hodeidah)	F	30	.44*	.51**
		M	15	.38	.29
2	Arts (Ibb)	F	80	.68**	.57**
		M	86	.64**	.31**
3	Computer Science and Engineering (Hod)	F	31	.27	.58**
		M	65	.27*	.52**
4	Dentistry (Hodeidah)	F	50	.36*	.54**
		M	17	.26	.39
5	Dentistry (Ibb)	F	42	.47**	.40**
		M	26	.05	.30
6	Education (Hodeidah)	F	79	.42**	.48**
		M	28	.37	.32
7	Engineering and Architecture (Ibb)	F	20	.25	.30
		M	31	.24	.13
8	Medical Science (Hodeidah)	F	91	.27**	.76**
		M	96	.22*	.66**
9	Sciences (Ibb)	F	38	.63**	.25
		M	56	.45**	-.08

* < .05, ** < .01. HSGPA = High School Grade Point Average, CETS = College Entrance Test Scores

Table 12. Correlation Coefficients between Predictor Variables and First-year College GPA by High School Location

No.	College	Location	N	HSGPA (r_{xy})	CETS (r_{xy})
1	Arts (Hodeidah)	R	21	.50*	.24
		U	24	.29	.57**
2	Arts (Ibb)	R	66	.67**	.41**
		U	100	.66**	.47**
3	Computer Science and Engineering (Hod)	R	38	.21	.56**
		U	58	.30*	.53**
4	Dentistry (Hodeidah)	R	11	.23	.23
		U	56	.35**	.50**
5	Dentistry (Ibb)	R	25	-.001	.35
		U	43	.54**	.44**
6	Education (Hodeidah)	R	44	.39**	.37*
		U	63	.42**	.54**
7	Engineering and Architecture (Ibb)	R	23	.21	.27
		U	28	.22	.19
8	Medical Science (Hodeidah)	R	63	.38**	.60**
		U	124	.16	.77**
9	Sciences (Ibb)	R	30	.59**	-.19
		U	64	.47**	.18

* < .05, ** < .01. HSGPA = High School Grade Point Average, CETS = College Entrance Test Scores

A number of multiple regression analyses were performed to answer research questions two through five. The second research question asked if high school grade point average and college entrance test scores were significant predictors of first-year college grade point average and whether the addition of college entrance test scores to high school GPA improved the prediction power of college performance. Nine colleges use college entrance test scores along with high school GPA to select their students at both universities. Hierarchical regression analyses were implemented to examine the predictive validity of high school GPA and college entrance test scores for each of the nine colleges separately. High school GPA and college entrance test scores were entered in blocks, Model 1 and Model 2, respectively. The change in R^2 (coefficient of determination) was appraised to see if college entrance test scores provided incremental information for the prediction of first-year college GPA.

In Model 1, first-year college grade point average was regressed on high school grade point average for each college alone. Table 13 presents the regression coefficients and the proportion of variance (R^2) in first-year college GPA accounted for by high school GPA. *F*-test results showed that high school GPA was a statistically significant predictor at all colleges except for the College of Engineering and Architecture, $F(1, 49) = 2.37$, $p = 0.130$ at Ibb University. Although high school GPA appeared to be a statistically significant predictor at all hard-science colleges, it accounted for less than 12% of the total variance in first-year college GPA at these colleges (except for the College of Sciences which accounted for 26% of the total variance in first-year college GPA). High school GPA made its highest statistically significant contribution at the College of Arts at Ibb University, $F(1, 164) = 126.79$, $p < .01$, $R^2 = .44$ and its lowest statistically significant contribution was at the College of Medical Science at Hodeidah

University, $F(1, 185) = 11.664$, $p < .01$, $R^2 = .06$. The results also showed that College of Arts at Ibb University significantly explained 44% of the variance in first-year college GPA whereas College of Arts at Hodeidah University significantly explained 14% of the total variance in first-year college GPA. This may be due to the differences in sample sizes at these colleges, $N = 166$ and $N = 45$, respectively.

Regression coefficients are another indicator of the relative strength of the predictor variables. They may be interpreted as number of points that a dependent variable (first-year college GPA) changes for a unit change in a predictor variable (high school GPA), all other factors held constant. For example, a student's first-year college GPA increases by .36 points for every unit increase in high school GPA at the College of Arts at Hodeidah University.

Table 13. Regression Coefficients and Percent of Variance in First-year College GPA Explained by High School GPA in Model 1

No.	College	N	B_{HSGPS}	R^2
1	Arts (Hodeidah)	45	.360*	14 %
2	Arts (Ibb)	166	1.324*	44 %
3	Computer Science and Engineering (Hodeidah)	96	.425*	7 %
4	Dentistry (Hodeidah)	67	.413*	11 %
5	Dentistry (Ibb)	68	1.923*	12 %
6	Education (Hodeidah)	107	.438*	16 %
7	Engineering and Architecture (Ibb)	51	.624	5 %
8	Medical Science (Hodeidah)	187	.340*	6 %
9	Sciences (Ibb)	94	1.325*	26 %

* = Statistically significant at .05 level. HSGPA = High School Grade Point Average.

The second part of the research question was intended to examine if the addition of college entrance test scores increased the predictive power of college performance. Model 1 (Table 13) shows that high school GPA was a statistically significant predictor of first-year college GPA at all colleges except the College of Engineering and Architecture at Ibb University. Table 14 presents the regression coefficients for high school GPA and college entrance test scores in Model 2 where the college entrance test scores variable was entered into the prediction equation along with high school GPA.

The addition of college entrance test scores as a predictor in Model 2 made a statistically significant contribution to R^2 at all colleges, except the College of Engineering and Architecture and the College of Sciences at Ibb University. The table also presents the percentage of explained variance (R^2) in first-year college GPA that was accounted for by the two predictor variables, high school GPA and college entrance test scores, and the change in R^2 in Model 2 after college entrance test scores were added to Model 1 (high school GPA). The results indicated that the addition of college entrance test scores was important because it accounted for a considerable amount of additional variance in first-year college GPA when added to the prediction equation. However, the importance was not observed across all colleges. Interestingly, the relative importance of including college entrance test scores was more pronounced at the colleges for the Hodeidah University and relatively unimportant at the colleges for the Ibb University.

Table 14. Regression Coefficients and Percent of Variance in First-year College GPA Explained by High School GPA and College Entrance Test Scores in Model 2

No.	College	N	B_{HSGPA}	B_{CETS}	R^2	R^2 Change
1	Arts (Hodeidah)	45	.340*	.261*	34%	20 %
2	Arts (Ibb)	166	1.163*	.173*	50%	6 %
3	Computer Sci. and Engin. (Hodeidah)	96	.275	.450*	31%	24 %
4	Dentistry (Hodeidah)	67	.323*	.221*	27%	16 %
5	Dentistry (Ibb)	68	1.499*	.213*	24%	12 %
6	Education (Hodeidah)	107	.224	.364*	23%	7 %
7	Engineering and Architecture (Ibb)	51	.473	.075	7%	2 %
8	Medical Science (Hodeidah)	187	.185*	.651*	52%	46 %
9	Sciences (Ibb)	94	1.322*	.015	26%	0 %

* = Statistically significant at .05 level. HSGPA = High School Grade Point Average, CETS = College Entrance Test Scores

4.4 DIFFERENTIAL PREDICTION BY GENDER AND HIGH SCHOOL LOCATION

The third research question asked if high school grade point average and college entrance test scores are differentially predictive of academic performance in Yemeni colleges across gender (male/female) and high school location (rural/urban). The Gulliksen and Wilks (1950) procedure was used to test for whether different prediction equations should be used with different subgroups of students. This procedure involves a test of equality of standard errors of estimate,

followed by a test of equality of regression slopes, and then a test of equality of regression intercepts across the different subgroups. Standard errors of estimate, regression slopes, and regression intercepts for males and females and for rural and urban students at each college were obtained and discussed below.

4.4.1 Differential Prediction by Gender

Gulliksen and Wilks test results for male and female students on high school grade point average at each college are reported in Table 15 and on college entrance test score in Table 16. Table 17 provides a summary of the findings of the Gulliksen and Wilks tests. Standard errors of estimate were found to be significantly different for only the College of Arts at Hodeidah University on both high school GPA and college entrance test scores, $F(28,13) = 3.892$, $p = .006$ and $F(28,13) = 3.32$, $p = .013$, respectively. There were significant differences on high school GPA, however, between male and female students on regression slopes for all except one college, the College of Engineering and Architecture, $F(2,47) = 1.506$, $p = .232$, at Ibb University, and there were significant differences on college entrance test scores for all except two colleges at Ibb University, the College of Engineering and Architecture, $F(2,47) = 1.263$, $p = .292$, and College of Sciences, $F(2,90) = 1.460$, $p = .237$. For the regression intercepts, there were no significant differences on high school GPA. There was only one significant difference on college entrance test scores at the College of Arts at Ibb University, $F(1,162) = 5.223$, $p = .024$. In summary, the results summarized in Table 17 show that significant differences in regression slopes were found for male versus female students at all colleges except one on high school GPA and two on college entrance test scores. This indicates that using common regression equations derived from

a pooled sample of male and female students at those colleges to predict students' academic performance may result in predictive bias against certain subgroups.

Table 15. Standard Errors of Estimate, Regression Slopes, and Regression Intercepts for High School GPA by Gender

No	College	Female				Male			
		N	SEE	B_1	B_0	N	SEE	B_1	B_0
1	Arts (Hodeidah)	30	4.81	.43	39.26	15	2.44	.42	42.92
2	Arts (Ibb)	80	6.39	1.32	-25.13	86	6.20	1.31	-26.67
3	Computer Sci. and Engin. (Hod)	31	6.43	.37	41.67	65	6.78	.45	33.38
4	Dentistry (Hodeidah)	50	3.94	.46	40.84	17	3.65	.29	56.27
5	Dentistry (Ibb)	42	6.97	2.41	-133.3	26	6.55	.29	46.18
6	Education (Hodeidah)	79	5.34	.51	33.32	28	4.15	.29	52.50
7	Engin. and Archit. (Ibb)	20	5.40	1.02	-8.86	31	4.27	.57	26.49
8	Medical Science (Hodeidah)	91	5.80	.40	41.89	96	5.90	.30	50.53
9	Sciences (Ibb)	38	6.48	1.95	-74.92	56	6.83	1.06	-8.97

SEE = Standard Error of Estimate, B_1 = Regression Slope, B_0 = Regression Intercept.

Table 16. Standard Errors of Estimate, Regression Slopes, and Regression Intercepts for College Entrance Test Scores by Gender

No.	College	Female				Male			
		N	SEE	B_1	B_0	N	SEE	B_1	B_0
1	Arts (Hodeidah)	30	4.61	.31	52.29	15	2.53	.13	65.35
2	Arts (Ibb)	80	7.16	.41	41.89	86	7.70	.18	59.16
3	Computer Sci. and Engin. (Hod)	31	5.43	.55	28.86	65	6.02	.45	35.10
4	Dentistry (Hodeidah)	50	3.54	.34	55.14	17	3.47	.17	69.64
5	Dentistry (Ibb)	42	7.21	.28	55.28	26	6.26	.15	61.22
6	Education (Hodeidah)	79	5.18	.53	27.21	28	4.23	.32	46.63
7	Engin. and Archit. (Ibb)	20	5.33	.16	62.50	31	4.36	.06	69.49
8	Medical Science (Hodeidah)	91	3.91	.71	22.61	96	4.55	.63	28.92
9	Sciences (Ibb)	38	8.03	.22	57.03	56	7.63	-.05	75.31

SEE = Standard Error of Estimate, B_1 = Regression Slope, B_0 = Regression Intercept.

Table 17. A Summary of the Gulliksen and Wilks Regression Tests by Gender

No.	College	HSGPA			CETS		
		SEE	B_1	B_0	SEE	B_1	B_0
1	Arts (Hodeidah)	*	*	ns	*	*	ns
2	Arts (Ibb)	ns	*	ns	ns	*	*
3	Computer Science and Engineering (Hod)	ns	*	ns	ns	*	ns
4	Dentistry (Hodeidah)	ns	*	ns	ns	*	ns
5	Dentistry (Ibb)	ns	*	ns	ns	*	ns
6	Education (Hodeidah)	ns	*	ns	ns	*	ns
7	Engineering and Architecture (Ibb)	ns	ns	ns	ns	ns	ns
8	Medical Science (Hodeidah)	ns	*	ns	ns	*	ns
9	Sciences (Ibb)	ns	*	ns	ns	ns	ns

SEE = Standard Error of Estimate, B_1 = Regression Slope, B_0 = Regression Intercept, HSGPA = High School GPA, CETS = College Entrance Test Scores, * = Significant, ns = Not Significant.

4.4.2 Differential Prediction by High School Location

Differential prediction analysis using Gulliksen and Wilks procedures was also conducted for students who came from rural and urban high schools for each college. The results are reported in Tables 18 through 20. Standard errors of estimate were found to be not significantly different for any of the 18 comparisons on each of the predictor variables except at the College of Medical Science at Hodeidah University on college entrance test scores, $F(61,122) = 1.69$, $p = .007$. There were significant differences between rural and urban students in regression slopes for all except one college, the College of Engineering and Architecture at Ibb University on high school GPA, $F(2,47) = 1.150$, $p = .325$, and significant differences on college entrance test scores for all except two colleges at Ibb University, the College of Engineering and Architecture, $F(2,47) = 1.348$, $p = .270$, and College of Sciences, $F(2,90) = 1.553$, $p = .217$. For the regression intercepts, there was only one significant difference on high school GPA at the College of Dentistry at Ibb University, $F(1,64) = 4.841$, $p = .03$, and one significant difference on college entrance test scores at the College of Medical Science at Hodeidah University, $F(1,183) = 3.980$, $p = .048$. In summary, as for the results based on gender subgroups, the results summarized in Table 20 indicate predictive bias across common regression equations derived from a pooled sample of urban and rural high school students.

Table 18. Standard Errors of Estimate, Regression Slopes, and Regression Intercepts for High School GPA by High School Location

No.	College	Rural				Urban			
		N	SEE	B_1	B_0	N	SEE	B_1	B_0
1	Arts (Hodeidah)	21	3.92	.57	28.26	24	4.31	.23	56.69
2	Arts (Ibb)	66	6.50	1.40	-32.53	100	6.31	1.28	-23.23
3	Computer Sci. and Engin. (Hod)	38	7.00	.35	43.03	58	6.49	.47	31.87
4	Dentistry (Hodeidah)	11	4.37	.25	58.18	56	3.75	.45	42.32
5	Dentistry (Ibb)	25	6.64	-.01	76.69	43	6.79	2.99	-186.26
6	Education (Hodeidah)	44	4.98	.43	40.79	63	5.18	.47	36.67
7	Engin. and Architecture (Ibb)	23	4.88	.67	18.87	28	4.72	.60	24.73
8	Medical Science (Hodeidah)	63	5.72	.50	32.59	124	5.88	.24	55.87
9	Sciences (Ibb)	30	6.88	1.77	-63.07	64	6.72	1.16	-15.44

SEE = Standard Error of Estimate, B_1 = Regression Slope, B_0 = Regression Intercept.

Table 19. Standard Errors of Estimate, Regression Slopes, and Regression Intercepts for College Entrance Test Scores by High School Location

No.	College	Rural				Urban			
		N	SEE	B_1	B_0	N	SEE	B_1	B_0
1	Arts (Hodeidah)	21	4.40	.18	60.10	24	3.69	.27	55.75
2	Arts (Ibb)	66	7.98	.26	53.41	100	7.39	.30	50.38
3	Computer Sci. and Engin. (Hod)	38	5.96	.53	30.41	58	5.78	.46	34.80
4	Dentistry (Hodeidah)	11	4.37	.12	71.33	56	3.48	.27	61.02
5	Dentistry (Ibb)	25	6.22	.18	62.85	43	7.25	.28	52.58
6	Education (Hodeidah)	44	5.01	.44	36.58	63	4.81	.60	20.66
7	Engin. and Architecture (Ibb)	23	4.81	.14	62.77	28	4.75	.09	67.30
8	Medical Science (Hodeidah)	63	4.93	.55	35.12	124	3.79*	.75	19.91
9	Sciences (Ibb)	30	8.38	-.15	82.43	64	7.50	.13	63.70

SEE = Standard Error of Estimate, B_1 = Regression Slope, B_0 = Regression Intercept.

Table 20. A Summary of the Gulliksen and Wilks Regression Tests by High School Location

No.	College	HSGPA			CETS		
		SEE	B_1	B_0	SEE	B_1	B_0
1	Arts (Hodeidah)	ns	*	ns	ns	*	ns
2	Arts (Ibb)	ns	*	ns	ns	*	ns
3	Computer Science and Engineering (Hod)	ns	*	ns	ns	*	ns
4	Dentistry (Hodeidah)	ns	*	ns	ns	*	ns
5	Dentistry (Ibb)	ns	*	*	ns	*	ns
6	Education (Hodeidah)	ns	*	ns	ns	*	ns
7	Engineering and Architecture (Ibb)	ns	ns	ns	ns	ns	ns
8	Medical Science (Hodeidah)	ns	*	ns	*	*	*
9	Sciences (Ibb)	ns	*	ns	ns	ns	ns

SEE = Standard Error of Estimate, B_1 = Regression Slope, B_0 = Regression Intercept, HSGPA = High School GPA, CETS = College Entrance Test Scores, * = Significant, ns = Not Significant.

4.5 PREDICTIVE VALIDITY AS MEASURED BY FOUR-YEAR CUMULATIVE GRADE POINT AVERAGE

The predictive validity of high school grade point average and college entrance test scores as measured by four-year cumulative grade point average was also examined for each college at Hodeidah University and Ibb University separately. The correlation coefficients (r_{xy}) between high school GPA and college entrance test scores with four-year cumulative grade point average are presented in Table 21. The results showed that there were significant low-to-moderate correlations between high school GPA and four-year cumulative GPA at all colleges except three, the College of Dentistry at Hodeidah University, r_{xy} = 0.17, p = .33, and College of Dentistry, r_{xy} = 0.18, p = .20, and College of Engineering and Architecture, r_{xy} = 0.19, p = .21, at Ibb University.

The results also showed that there were significant moderate-to-high correlations between college entrance test scores and four-year cumulative GPA at all colleges except two, the College of Dentistry, r_{xy} = 0.21, p = .14, and College of Sciences, r_{xy} = 0.07, p = .52, at Ibb University. The strongest significant relationship between high school GPA and four-year cumulative GPA was at the College of Arts at Ibb University, r_{xy} = 0.60, p < .01; while the lowest significant relationship was at the College of Medical Science at Hodeidah University, r_{xy} = 0.19, p < .05. Furthermore, the strongest significant relationship between college entrance test scores and four-year cumulative GPA was at the College of Computer Science and Engineering at Hodeidah University, r_{xy} = 0.82, p < .01; while the lowest significant relationship was at the College of Engineering and Architecture at Ibb University, r_{xy} = 0.34, p < .05. The results revealed, in

general, that the magnitude of the relationship between college entrance test scores and four-year cumulative GPA was higher than the relationship between high school GPA and four-year cumulative GPA.

In comparison with first-year college GPA, the relationship between high school GPA and four-year cumulative GPA was similar to the relationship between high school GPA and first-year college GPA at most colleges. However, the correlations between college entrance test scores and four-year cumulative GPA were higher than the correlations between college entrance test scores and first-year college GPA at some colleges. Similarly, the magnitude of the relationship between the predictor variables and four-year cumulative GPA was higher for college entrance test scores than for high school GPA at most colleges.

The correlation coefficients of the relationship between high school GPA and college entrance test scores with four-year cumulative GPA at each college were also obtained by gender and high school location and are presented in Tables 22 and 23. The findings revealed that some colleges that had significant relationships between high school GPA and college entrance test scores with four-year cumulative GPA did not have significant relationships when analyzed by gender and high school location. For example, see the College of Engineering and Architecture at Ibb University when college entrance test scores were analyzed by gender and the College of Medical Science at Hodeidah University when high school GPA was analyzed by high school location. This is probably due to the decrease in the sample sizes and restrictions in range when examining subgroups.

Table 21. Correlation Coefficients between Predictor Variables and Four-year Cumulative GPA

No.	College	N	HSGPA (r_{xy})	# of Subtests	CETS (r_{xy})
1	Arts (Hodeidah)	42	.34*	1	.56**
2	Arts (Ibb)	157	.60**	1	.53**
3	Computer Science and Engineering (Hod)	84	.34**	3	.82**
4	Dentistry (Hodeidah)	36	.17	4	.71**
5	Dentistry (Ibb)	51	.18	4	.21
6	Education (Hodeidah)	89	.46**	1	.44**
7	Engineering and Architecture (Ibb)	45	.19	3	.34*
8	Medical Science (Hodeidah)	173	.19*	3	.64**
9	Sciences (Ibb)	87	.49**	3	.07

* < .05, ** < .01. HSGPA = High School GPA, CETS = College Entrance Test Scores.

Table 22. Correlation Coefficients between Predictor Variables and Four-year Cumulative GPA by Gender

No.	College	Gender	N	HSGPA (r_{xy})	CETS (r_{xy})
1	Arts (Hodeidah)	F	28	.39*	.56**
		M	14	.35	.66*
2	Arts (Ibb)	F	77	.60**	.67**
		M	80	.61**	.40**
3	Computer Science and Engineering (Hod)	F	23	.25	.97**
		M	61	.39**	.77**
4	Dentistry (Hodeidah)	F	28	.11	.61**
		M	8	.01	.93**
5	Dentistry (Ibb)	F	33	.17	.26
		M	18	.13	-.04
6	Education (Hodeidah)	F	68	.59**	.46**
		M	21	.12	.39
7	Engineering and Architecture (Ibb)	F	17	-.22	.30
		M	28	.32	.36
8	Medical Science (Hodeidah)	F	85	.21	.70**
		M	88	.17	.58**
9	Sciences (Ibb)	F	36	.53**	.23
		M	51	.48**	-.00

* < .05, ** < .01. HSGPA = High School GPA, CETS = College Entrance Test Scores.

Table 23. Correlation Coefficients between Predictor Variables and Four-year Cumulative GPA by High School Location

No.	College	Gender	N	HSGPA (r_{xy})	CETS (r_{xy})
1	Arts (Hodeidah)	R	19	.44	.47*
		U	23	.28	.60**
2	Arts (Ibb)	R	60	.65**	.46**
		U	97	.57**	.58**
3	Computer Science and Engineering (Hod)	R	31	.25	.76**
		U	53	.40**	.86**
4	Dentistry (Hodeidah)	R	3	-.07	.93
		U	33	.18	.69**
5	Dentistry (Ibb)	R	17	-.12	.19
		U	34	.34*	.26
6	Education (Hodeidah)	R	35	.61*	.52**
		U	54	.36**	.40**
7	Engineering and Architecture (Ibb)	R	19	.15	.31
		U	26	.19	.26
8	Medical Science (Hodeidah)	R	57	.26	.52**
		U	116	.15	.71**
9	Sciences (Ibb)	R	25	.44*	-.02
		U	62	.51**	.11

* < .05, ** < .01. HSGPA = High School GPA, CETS = College Entrance Test Scores.

The fourth research question was intended to examine how well high school grade point average and college entrance test scores predict students' long-term academic success. Hierarchical regression analysis was utilized to examine the predictive validity of high school GPA and college entrance test scores for each of the nine colleges separately. High school GPA and college entrance test scores were entered into the regression equation in blocks, Model 1and Model 2, respectively. The change in R^2 was also obtained to examine if college entrance test scores increased the predictive power of college success.

Four-year cumulative grade point average was first regressed on high school grade point average alone for each college separately in Model 1. *F*-test results showed that high school GPA was a statistically significant predictor at all colleges except three hard-science colleges, the College of Dentistry at Hodeidah University, $F(1, 34) = .966$, $p = 0.333$, and College of Dentistry, $F(1, 49) = 1.690$, $p = 0.200$ and College of Engineering and Architecture, $F(1, 43) = 1.616$, $p = 0.211$ at Ibb University. Table 24 shows the regression coefficients and the proportion of variance (R^2) in four-year cumulative GPA accounted for by high school GPA. High school GPA made its highest statistically significant contribution at the College of Arts at Ibb University, $F(1, 155) = 87.82$, $p < .01$, $R^2 = 36\%$, while its lowest significant contribution was at the College of Medical Science at Hodeidah University, $F(1, 171) = 6.277$, $p = 0.013$, $R^2 = 4\%$. These results are similar to that of first-year college GPA even though the amount of explained variance in four-year cumulative GPA are slightly lower than the amount of explained variance in first-year college GPA at most colleges. This is normal because the sample size was decreased due to students' dropouts. Like first-year college GPA findings, The College of Arts at Ibb University was still the college that significantly accounted for the highest amount of variance in

the criterion variable (four-year cumulative GPA) and the College of Medical Science at Hodeidah University was the one that significantly accounted for the lowest amount of variance.

Table 24. Regression Coefficients and the Percent of Variance in Four-year Cumulative GPA Explained by High School GPA in Model 1

No.	College	N	B_{HSGPA}	R^2
1	Arts (Hodeidah)	42	.414*	12 %
2	Arts (Ibb)	157	1.185*	36 %
3	Computer Science and Engineering (Hodeidah)	84	.578*	12 %
4	Dentistry (Hodeidah)	36	.393	3 %
5	Dentistry (Ibb)	51	.595	3 %
6	Education (Hodeidah)	89	.742*	21 %
7	Engineering and Architecture (Ibb)	45	.459	4 %
8	Medical Science (Hodeidah)	173	.270*	4 %
9	Sciences (Ibb)	87	1.306*	24 %

* = Statistically significant at .05 level. HSGPA = High School GPA.

Next, the predictive validity of the addition of college entrance test scores to the prediction of four-year cumulative grade point average was investigated. It was hypothesized that college entrance test scores significantly predict academic performance after four years in college as measured by four-year cumulative GPA. The effect of high school GPA and college entrance test score variables on four-year cumulative GPA was examined using hierarchical regression analysis. The addition of college entrance test scores as a predictor in Model 2 made a statistically significant contribution to R^2. It increased the amount of explained variance in four-year cumulative GPA at all colleges, except two, the College of Dentistry, $F(2, 48) = 1.52$, $p = 0.229$, and College of Engineering and Architecture, $F(2, 42) = 2.86$, $p = 0.07$, at Ibb University. The results indicated that the addition of college entrance test scores had a significant contribution in the prediction of academic success over time. College entrance test scores accounted for a large amount of variance in the criterion variable, four-year cumulative GPA, when added to high school GPA. Table 25 presents the regression coefficients and the percentage of explained variance (R^2) in four-year cumulative GPA that was accounted for by the two predictor variables, high school GPA and college entrance test scores, and the change in R^2 in Model 2 where the college entrance test scores variable was added to the model with only high school GPA. As for the results based on first-year college GPA, college entrance test scores showed a significant incremental prediction only in the colleges at the Hodeidah University.

Table 25. Regression Coefficients and the Percent of Variance in Four-year Cumulative GPA Explained by High School GPA and College Entrance Test Scores in Model 2

No.	College	N	B_{HSGPA}	B_{CETS}	R^2	R^2 Change
1	Arts (Hodeidah)	42	.367*	.403*	41 %	29 %
2	Arts (Ibb)	157	.952*	.238*	50 %	14 %
3	Computer Sci. and Engin. (Hodeidah)	84	.314*	.736*	70 %	58 %
4	Dentistry (Hodeidah)	36	.227	.707*	51 %	48 %
5	Dentistry (Ibb)	51	.419	.057	6%	3 %
6	Education (Hodeidah)	89	.502*	.413*	26%	5 %
7	Engineering and Architecture (Ibb)	45	.155	.134	12%	8 %
8	Medical Science (Hodeidah)	173	.185*	.644*	43%	39 %
9	Sciences (Ibb)	87	1.302*	.041	25%	1 %

* = Statistically significant at .05 level. HSGPA = High School GPA, CETS = College Entrance Test Scores.

4.6 PREDICTIVE VALIDITY OF STUDENTS PERSISTENCE TO GRADUATION

The last research question investigated if high school grade point average and college entrance test scores were significant predictors of students' persistence as measured by whether a student graduated or not within the norm-time of four years of study in college. A logistic regression analysis was used to answer this question. Out of the pooled sample size of 881, only 117 (13.28%) students did not graduate at the norm-time of four years in college. The frequency of students who passed and those who not pass for each college is presented in Table 26. Although the College of Medical Science at Hodeidah University had the highest number of enrolled students (187), only 14 (7%) did not graduate at the norm-time of four years of college. The College of Arts at Ibb University had the second highest number of enrolled students (166) with the lowest number of students who did not graduate: 9 (5%).

Table 26. Frequency of Students who Passed/Did not Pass at each College

No.	College	Passed		Not Passed	
		N	%	N	%
1	Arts (Hodeidah)	42	93	3	7
2	Arts (Ibb)	157	95	9	5
3	Computer Science and Engineering (Hodeidah)	84	87	12	13
4	Dentistry (Hodeidah)	36	54	31	46
5	Dentistry (Ibb)	51	75	17	25
6	Education (Hodeidah)	89	82	18	18
7	Engineering and Architecture (Ibb)	45	88	6	12
8	Medical Science (Hodeidah)	173	93	14	7
9	Sciences (Ibb)	87	93	7	7

The results of logistic analysis revealed that high school grade point average was not a significant predictor of students' persistence at any of the participating colleges. It did not have a significant effect on the probability of a student to graduate from college in a four-year timespan at any of the colleges in the study (Table 27). When college entrance test scores were added to high school GPA, it significantly improved the predictive validity at four hard-science colleges. These colleges were the College of Dentistry at Hodeidah University, χ^2 (2) = 14.688, $p < .01$,

Negelkerke R^2 = .26; and College of Engineering and Architecture, χ^2 (2) = 13.641, p < .01, Negelkerke R^2 = .46; College of Medical Science, χ^2 (2) = 15.478, p < .01, Negelkerke R^2 = .19; and College of Sciences, χ^2 (2) = 9.364, p < .01, Negelkerke R^2 = .23 at Ibb University.

Table 28 presents the regression coefficients, odds ratios, and the percent of the explained variance in students' persistence that was accounted for by the two predictors, high school GPA and college entrance test scores, and the change in Negelkerke R^2 in Model 2 where college entrance test scores were added to the model with high school GPA. The odds ratios are often used to quantify the effect of significant predictor variables (high school GPA and college entrance test scores) on the criterion variable (persistence). The table shows that two colleges have significant odds ratio less than one and two have significant odds ratio greater than one.

Odds less than one imply that the change in odds that a student will graduate from college in the norm-time of four years declines when college entrance test scores are increased by one point. In other words, the higher the college entrance test scores, the less likely it is that a student would graduate in the norm-time of four years in college. For example, the change in odds that a student will graduate in the norm-time of four years at the College of Sciences at Ibb University decreases by .882 for each point increase on the college entrance test scores. In contrast, odds greater than one imply that the change in odds that a student will graduate from college in the norm-time of four years increases when college entrance test scores are increased by one point. For example, at the College of Dentistry at Hodeidah University, if the college entrance test scores are increased by one point, the change in odds that a student will graduate from this college in the norm-time of four years will increase by a factor of 1.148, when high school GPA is controlled for.

Table 27. Regression Coefficients, Odds Ratios, and Percent of Variance in Persistence Variable Explained by High School GPA in Model 1

No.	College	N	*B*	e^B	Negelkerke R^2
1	Arts (Hodeidah)	45	-.041	.960	1 %
2	Arts (Ibb)	166	.076	1.079	1 %
3	Computer Sci. and Engin. (Hodeidah)	96	.099	1.104	3 %
4	Dentistry (Hodeidah)	67	.098	1.103	3 %
5	Dentistry (Ibb)	68	.102	1.107	1 %
6	Education (Hodeidah)	107	.003	1.003	0 %
7	Engineering and Architecture (Ibb)	51	.113	1.120	1 %
8	Medical Science (Hodeidah)	187	.088	1.092	2 %
9	Sciences (Ibb)	94	-.085	.919	1 %

* = *Significant at .05level.*

Table 28. Regression Coefficients, Odds Ratios, and the Percent of Variance in Four-year Cumulative GPA Explained by High School GPA and College Entrance Test Scores in Model 2

No.	College	HSGPA		CETS		Negelkerke R^2	R^2 Change
		B	e^B	*B*	e^B		
1	Arts (Hodeidah)	-.035	.970	.252	1.286	23 %	22 %
2	Arts (Ibb)	.078	1.081	-.002	.998	1 %	0 %
3	Computer Sci. and Engin. (Hod)	.110	1.116	-.030	.971	4 %	1 %
4	Dentistry (Hodeidah)	.058	1.060	.138	1.148*	26 %	23 %
5	Dentistry (Ibb)	.068	1.071	.018	1.018	2 %	1 %
6	Education (Hodeidah)	.023	1.024	-.034	.966	1 %	1 %
7	Engin. and Architecture (Ibb)	.682	1.978	-.235	.791*	46 %	45 %
8	Medical Science (Hodeidah)	.042	1.043	.162	1.176*	19 %	17 %
9	Sciences (Ibb)	-.062	.939	-.126	.882*	23 %	22 %

* = Statistically significant at .05 level. HSGPA = High School GPA, CETS = College Entrance Test Scores.

5.0 DISCUSSION

This chapter provides a summary and interpretation of the findings. Conclusions and implications derived from the findings are also reported in this chapter. The limitations of the study are presented, and recommendations for further research are also suggested.

5.1 SUMMARY

Bachman and Palmer (1996) argued that one of the most important considerations in designing and developing an assessment is the use for which it is intended. They explained that the usefulness of any assessment system is judged by its effectiveness in achieving its purpose. In this study, the focus was on the admission system for two universities in Yemen. High school grade point average and college entrance test scores are two criteria commonly used for admission in Yemeni colleges. High school GPA provides admission personnel with a sense of students' achievements in high school, and college entrance test scores provide them with an index of students' potential to perform well in college. Admission personnel need to have the most proper criteria to ensure valid and fair admission decisions. Willingham, Lewis, Morgan, and Ramist (1990) indicated that any measure that is commonly used for selection has direct and indirect effects throughout the educational system and that this measure needs to validate or justify that use. The admission decision can make a critical impact on students' futures and on

the quality of the output of the education system. Therefore, the question of whether high school GPA and college entrance test scores used as admission criteria are sufficient predictors of future academic success should always be validated to ensure fair admission decisions. There are many types of validity evidence; predictive validity is one of them. A measure is said to have predictive validity to the extent that it predicts students' future academic success.

The purpose of this study was to investigate the predictive validity of both high school grade point average and college entrance test scores used as admission criteria to postsecondary institutions in Yemen. The study also examined if adding college entrance test scores to high school GPA would increase the power to predict first-year college grade point average and four-year cumulative grade point average. The differential predictive validity of high school GPA and college entrance test scores was examined across gender and high school location. The relationship between students' persistence in the four years of college and the predictor variables, high school GPA and college entrance test scores, was studied as well. Before the predictive validity analyses were performed, the study examined the reliability of college entrance test scores using internal consistency measure. The college entrance tests consisted of different subject areas. It was crucial to examine if the individual as well as the composite entrance tests were truly reliable because college entrance tests' reliability had not been examined previously.

Data for 881 students were obtained from the students' records at two universities in Yemen, Hodeidah University and Ibb University. A number of statistical techniques were utilized to help answer the research questions, namely internal consistency, hierarchical multiple and logistic regression analyses, and tests of equality of standard errors of estimates, regression slopes, and regression intercepts. Descriptive analyses were also included to provide general

information about students and predictor and criterion variables. Analyses were conducted separately for each college to examine the predictive validity of high school GPA and college entrance test scores to predict short- and long-term academic performance. Analyses were performed separately because entrance tests are college-dependent. Differential prediction analyses were also obtained across gender and high school location for each college.

5.2 INTERPRETATION OF THE FINDINGS

The descriptive analyses for the predictor variables (high school GPA and college entrance test scores) and the criterion variables (first-year college GPA and cumulative GPA) were first obtained. Female students' mean scores were found to be higher than male students' mean scores on the predictor variables as well as on the criterion variables. Similarly, urban school students' mean scores were found to be higher than rural school students' mean scores on high school GPA and on the criterion variables, first-year college GPA and four-year cumulative GPA. Rural school students' mean score, however, were higher than urban school students' on college entrance test scores. Differential predictive validity of high school GPA and college entrance test scores across the different colleges is discussed in a later section. The findings also showed that students majoring in hard-science colleges (Computer, Dentistry, Medical Science, and Engineering) had higher mean scores than those in other majors. This was due to the fact that these colleges require a high school GPA of 80% and higher.

Literature consistently revealed that male and female students differ in their performance on various high school subject areas, standardized tests, and testing programs and that female students generally outperform male students (Wilder and Powell, 1989; Willingham and Cole, 1997; Hyde, Fennema, and Lamon, 1990; Azen, Bronner, & Gafni, 2002). In this study female students in Yemen also outperformed male students on high school GPA, college entrance test scores, first-year college GPA, and four-year cumulative GPA. The findings also showed that urban school students' achievements were found to be higher than rural school students. These results are different from Stanley, Comello, Edwards, and Marquart's (2008) findings. They found that there was no significant difference in the achievement of students in rural and urban areas. Grissmer, Flanagan, Kawatka, and Williamson (2000) in their regression study of students' achievements also concluded that location (rural, urban, or suburban) was not found to be a statistically significant predictor.

Question 1: What is the reliability, as measured by internal consistency, of the college entrance test scores used for admission into Yemeni universities?

The reliability coefficients of the college entrance test scores used as predictors of students' academic success were examined using Cronbach's alpha and stratified alpha coefficient analyses. At Hodeidah University, the reliability coefficient estimates ranged from .77 to .94. At Ibb University, the reliability coefficient estimates ranged from .79 to .92. The reliability coefficients of the composite tests used by different colleges were also examined. At Hodeidah University, the reliability coefficient estimates ranged from .85 to .88. At Ibb University, the reliability coefficient estimates ranged from .81 to .89. Thus, the results obtained from this study

showed that the reliability coefficient estimates of college entrance test scores were in the acceptable range. Reliability coefficients must approximate or exceed .80 in magnitude for any measure to be considered minimally reliable; coefficients must be .90 or above to be considered most desirable (Aiken, 1997; Helmstadter, 1964; Nunnally, 1978; Salvia & Ysseldyke, 1995).

Question 2: Are high school grade point average and college entrance test scores significant predictors of first-year college grade point average for each of the different colleges at Yemeni universities? Does the addition of college entrance test scores enhance the prediction of college performance? Are the results comparable to what has been found in the United States?

Multiple regression analyses were conducted to answer this research question. The significance of high school grade point average was examined first for each college separately. In a subsequent regression model, college entrance test scores were entered into the prediction equation to investigate their incremental effect on the prediction power.

The results revealed that high school GPA was a statistically significant predictor at all colleges except for the College of Engineering and Architecture at Ibb University. The addition of college entrance test scores to high school GPA improved the prediction power especially at hard-science colleges. In fact, college entrance test scores accounted for more variance in first-year college GPA than high school GPA did at most colleges. However, the relative importance of adding college entrance test scores was observed at the colleges for Hodeidah University than Ibb University colleges.

The findings of this study correspond with many studies which found that high school GPA and college entrance test scores are generally significant predictors of students' academic performance during their undergraduate studies in college (Astin, Korn, & Green, 1987; Noble, 1991; Moffat, 1993; Bridgeman, McCamley-Jenkins, & Ervin, 2000; Snyder, Hackett, Stewart, & Smith, 2003; Kim, 2002; Kuncel, Hezlett, & Ones, 2004; Ramist, Lewis, & McCamley-Jenkins, 1994; Waugh, Micceri, & Takalkar, 1994; Wolfe & Johnson, 1995; Kuncel et al., 2005; Kuncel, Credé, & Thomas, 2007). Mathiasen (1984) reviewed more than 60 studies and concluded that high school GPA and standardized entrance test scores are the best predictors of college performance when predicting first-year college GPA. Similar to the findings of Ramist, Lewis, & McCamley-Jenkins's (1994) comprehensive study, the association between high school GPA and first-year college GPA showed a moderate correlation of .39 when averaged across the participating colleges. Furthermore, when the correlation coefficients between college entrance test scores and first-year college GPA were also averaged across the colleges, the results yielded a correlation of .49.

Like many studies, the prediction power of college success was enhanced by adding college entrance test scores to high school GPA (Camara & Echternacht, 2000). Using high school GPA and college entrance test scores could help admissions personnel to make more accurate predictions and more appropriate admission decisions than using high school GPA alone (Eimers & Pike, 1997; Willingham, 1985; Mouw & Khanna, 1993). However, the results in this study showed that college entrance test scores explained more variance in first-year college GPA than high school GPA at many of the colleges in the study. These findings are different from most of the previous research which found that high school GPA explains more of the total variance in first-year college GPA than standardized entrance tests like the SAT or ACT

(Munro, 1981; Zheng et al., 2002; Hoffman, 2002; Cowen & Fiori, 1991; Elert, 1992; Myers & Pyles, 1992; Moffatt, 1993; Ramist, Lewis, and McCamley-Jenkins, 1993; Camara & Echternacht, 2000; Hu, 2002; Willingham, 1985; Zwick & Sklar, 2005).

Many researchers have further discussed that high school GPA more accurately predicts academic success in colleges than standardized tests or any other factor (Munro, 1981; Lawlor, Richman, & Richman, 1997; Peltier, Laden, & Martranga, 1999; Snyder, Hackett, Stewart, & Smith, 2003; Camara & Echternacht, 2000; Tross et al, 2000; Fleming & Garcia, 1998; Fleming, 2002; Hoffman, 2002; Zheng et al., 2002; Gose, 1994) and that standardized test scores add little information to prediction equations beyond high school GPA (Cowen & Fiori, 1991; Moffatt, 1993; Myers & Pyles, 1992). These conclusions require us to look back and investigate other types of validity evidence (i.e. content and internal consistency) for high school tests. It is necessary to find out why it is not the case in the Yemeni context that high school GPA is consistently the best predictor of first-year college GPA and that standardized test scores do add a statistically significant increment to the prediction. Given the many criticisms of high school tests in Yemen mentioned in the introduction of this study, the use of high school GPA is generally doubted as a measure to predict academic success and admit students to colleges.

Question 3: Do high school grade point average and college entrance test scores have differential prediction when used to predict academic performance in Yemeni colleges across gender and high school location? Are the results comparable to what has been found in the United States?

A test is considered biased if the prediction by its scores is not equivalent for different subgroups (Johnson, Carter, Davison, and Oliver, 2001). The Gulliksen and Wilks (1950) procedure was used in this study to test for the differential prediction by high school grade point average and college entrance test scores. The procedure involves a test of standard errors of estimate, a test of equality of the regression slopes, and a test of equality of the regression intercepts. For each college, if a significant difference was found for one of the comparisons, then it would be concluded that differential prediction occurs.

The findings for male and female students for each of the participating colleges showed that standard errors of estimate were not significantly different for all except one college, the College of Arts at Hodeidah University, on both predictor variables. There were significant differences between male and female students on regression slopes for all except one college (the College of Engineering and Architecture at Ibb University) on high school GPA and significant differences on college entrance test scores for all except two colleges (College of Engineering and Architecture and College of Sciences at Ibb University). For regression intercepts, there was only one significant difference on college entrance test scores at the College of Arts at Ibb University.

Differential prediction analyses for students coming from rural and urban high schools for each of the nine colleges were also conducted. The results indicated that standard errors of estimate were not significantly different for any of the 18 comparisons on each of the predictor variables except at the College of Medical Science at Hodeidah University on college entrance test scores. There were significant differences, however, between rural and urban students on regression slopes for all except one college (College of Engineering and Architecture at Ibb University) on high school GPA and significant differences on college entrance test scores for all except two colleges (College of Engineering and Architecture and College of Sciences at Ibb University). For regression intercepts, there was only one significant difference on high school GPA at the College of Dentistry at Ibb University and only one significant difference on college entrance test scores at the College of Medical Science at Hodeidah University.

In general, the results implied that using the common regression equation for male and female students and for those coming from rural and urban high schools may result in predictive bias against certain subgroups. In this study, predictive bias was found to favor female and urban students at most colleges. The presence of differential prediction indicates a lack of fairness in the tests' use. It is important for tests to be free from bias in selection decisions affecting minority groups (Equal Employment Opportunity Commission, 1978).

Regarding differential prediction of high school records and standardized admission tests in the United States, many studies examined the differences across race, ethnicity, gender, and socioeconomic status. It was believed that differential prediction is of great social and educational importance and is a real phenomenon that should be researched (Cole, 1972). Young (2001) reviewed and analyzed all the published studies of differential prediction of admission

tests during a period of more than 25 years (1974-2001). He found consistent results of differences in prediction among minority groups. College grades for minority groups were commonly overpredicted. For studies based on gender, he consistently found significant differences between male and female students and that female students were usually underpredicted (See Young, 2001 for a summary of the characteristics of each of the 37 studies, p. 19 and 20). According to Young, underprediction occurs when "the residuals are generally positive." Simply speaking, underprediction occurs when a subgroup predicted performance is below their actual performance.

Thomas (1972), in a study of differential prediction based on gender at 10 colleges, found that females' GPAs were underpredicted when a similar prediction equation was used to predict first-year GPA. Moreover, Sowa, Thomson, & Bennett (1989) strongly criticized the use of standardized tests in the admission process and asserted that these tests are gender- and race-biased and do not gauge the ability of certain student groups. In addition, some authors claimed that standardized tests do not predict achievement across gender and ethnic student groups (Bridgeman & Wendler, 1991; Moffatt, 1993).

Question 4: How well do high school grade point average and college entrance test scores predict students' long-term academic success in college? Are the results comparable to what has been found in the United States?

Researchers have criticized the narrowness of first-year college grade point average as a measure of college performance and have urged the use of other criteria, such as four-year cumulative grade point average and persistence (Geiser and Studley, 2003). This study examined the

predictive validity of high school grade point average and college entrance test scores as measured by four-year cumulative grade point average for each college separately at Hodeidah University and Ibb University. The results indicated that high school GPA was a statistically significant predictor at all colleges except three hard-science colleges (College of Dentistry at Hodeidah University and College of Dentistry and College of Engineering and Architecture at Ibb University). The addition of college entrance test scores as a predictor made a statistically significant incremental contribution to R^2 at all colleges, except two colleges (College of Dentistry and College of Engineering and Architecture at Ibb University). It increased the amount of explained variance in four-year cumulative GPA, especially at the hard-science colleges.

In general, the results indicated that high school GPA continued to significantly predict academic success over time and that the addition of college entrance test scores made a significant contribution to the prediction of academic success over time. The findings are consistent with Friedman and Mandel (2010) who concluded that high school GPA and standardized test scores significantly predict cumulative GPA. In fact, college entrance test scores accounted for a considerable amount of variance in the criterion variable, four-year cumulative GPA, when added to high school GPA, especially at hard-science colleges. The findings of four-year cumulative GPA were consistent with those for first-year college GPA. High school GPA and college entrance test scores were significant predictors of short- and long-term academic performance and the addition of college entranced test scores enhances the prediction power especially at hard-science colleges, in general, and the colleges at Hodeidah University, in particular.

The findings of this study, however, are different from a meta-analysis study by Elert (1992) who discussed that high school grades and college entrance test scores are good predictors of college performance and that the predictive power of entrance exams disappears after the first year in college. This study revealed that high school grades and college entrance test scores were significant predictors of academic success for first-year college GPA as well as four-year cumulative GPA.

Question 5: How well do high school grade point average and college entrance test scores predict students' persistence over time?

The results showed that out of the sample size of 881, only 117 (13.28%) students did not graduate in the norm-time of four years in college. The results also showed that high school grade point average was not a significant predictor of students' persistence at any of the participating colleges. This is different from results from Rosenbaum (2004) and Ishitani (2005) who found that high school grades have a significant effect on students' persistence in college. When college entrance test scores were added to high school GPA, it significantly improved the predictive validity at four hard-science colleges.

The findings in this study match those found by some previous studies in which high school GPA and standardized entrance scores were found to be not significant predictors of students' persistence (Lin, Yu, Chen, and Kauffman, 2011; Ting, 1998). Lohfink and Paulsen (2005) and Munro (1981) also found that entrance test scores do not have a direct tie to the retention of students by higher education institutions. The results in this study, however, are

different from results from Astin (1997) in which a large amount of variance in persistence could be explained by high school GPA and standardized test scores. It was also different from results from Friedman and Mandel (2010) in which students' high school GPA and SAT scores were significantly related to both cumulative GPA and retention after the first year.

Research on students' persistence has found that students do not finish college for a number of reasons (Ramist, 1981; Robbins, Lauver, Le, Davis, & Langley, 2004). These reasons can be related to academic factors, financial support, motivation, gender, ethnicity, socioeconomic status, and personal and institutional characteristics (Ramist, 1981). Bowen and Bok (1998) found that institutional characteristics were related to whether or not students persisted. Bowen & Bok (1998) and Cabrera, Nora, & Castañeda (1993) found that socioeconomic status (SES) was positively related to persistence in college. Tinto (1987) and Veltri, Banning, & Davis (2003) claimed that faculty-student interaction was a good predictor of student persistence. In Yemen none of the noncognitive factors are considered in the admission procedures. Evaluating which factors might help improve students' persistence would be extremely helpful to admission committees for selecting successful students.

5.3 LIMITATIONS

One of the limitations of this study is that the relationships between the predictor variables and the criterion variables were not corrected for restriction of range and criterion unreliability. Restrictions of range and criterion unreliability underestimate the true nature of the relationship between the predictor and the criterion variables. Correcting for restriction of range and criterion

unreliability is recommended because it is assumed that the corrected correlation between the predictor and the criterion variables is a better estimate of the population correlation than the observed correlation (Thorndike, 1949). However, correcting for restriction of range requires that the unrestricted variance of predictor variables be known. As this data was not available, this correction was not possible. In the context of predictive validity studies of academic success, corrections for criterion unreliability are not performed.

Another limitation is that the study was done with data from two public universities out of the eight universities in the country. Having more data from both public and private universities would provide additional information about the predictive validity of high school and admission tests in Yemen.

The third limitation is that students who did not graduate at the particular college were considered "failed to persist." The information of whether those students really failed to graduate in the norm-time or if they might have transferred to another institution to continue their education was not known.

The last limitation of the study was that students may not have had the opportunity to achieve high grades in high school because most of them have many other commitments in addition to school. In a country like Yemen, other commitments include working after school or taking care of their families as many get married at a young age. Such commitments might have lowered their high school grades and therefore had an impact on the results of the study.

5.4 CONCLUSION AND IMPLICATIONS

For any admission test the validity of the test for its intended purpose should be the primary consideration for the admission decision-makers. High school grade point average and college entrance test scores are commonly used in predicting future academic performance in higher education in Yemen. High school grades reflect students' performance in the twelfth grade in a variety of subjects. College entrance tests are based on subjects that university faculty regard as essential prerequisites for college level work and that research has shown to be highly correlated with college outcomes (Adelman, 1999).

This study is the first study that extensively evaluates the predictive validity of high school GPA and college entrance test scores for Yemeni universities. It is considered significant because it contributes to the body of knowledge concerning the predictive validity of high school GPA and college entrance test scores and may provide a baseline of information for future research.

The results revealed that high school grade point average and college entrance test scores are significant predictors of both short-term, as measured by first-year college GPA, and long-term academic success, as measured by four-year cumulative GPA. However, high school GPA demonstrated a poor predictive validity for both first-year college GPA and four-year cumulative GPA, and the addition of college entrance test scores to the prediction equation along with high school GPA enhanced the predictive power of college performance. In fact, college entrance test scores explained more variance in the criterion variables than high school GPA at the colleges for the Hodeidah University. The findings, however, are different from many studies that claim

standardized tests yield a small but statistically significant improvement in predicting short- and long-term college outcomes, beyond that which is provided by high school GPA. The poor predictive validity of high school GPA provides a clear need for admission decision-makers to comprehensively review the appropriateness of high school tests.

Another important conclusion is that differential prediction occurs across gender (male/female) and high school location (rural/urban). The findings of this study are not different from that which was found by Young (2001) in his review of all the published studies of differential prediction by gender during a period of more than 25 years (1974-2001). He observed consistently small to moderate differences in the accuracy of prediction equations. Young found consistent underprediction for female students.

Finally, high school grade point average was not found to be a significant predictor of students' persistence at any of the nine colleges. However, the prediction of college success was enhanced at four colleges when college entrance test scores were added to the prediction model. Research on predicting persistence from high school GPA and college entrance test scores claimed that persistence in college is influenced substantially by nonacademic factors and that high school GPA and college entrance test scores may not be good predictors of students' continuance in their studies (Hoffman & Lowitzki, 2005).

In conclusion, based on the findings of this study, admission personnel should give more weight to college entrance test scores because they have higher correlation coefficients with the criterion variables and have greater predictive power of college success than high school GPA. Differences among gender and high school location should also be taken into consideration

during the admission process to allow for more equal opportunities to all applicants and have fairer admission decisions.

5.5 RECOMMENDATIONS FOR FUTURE RESEARCH

To the best of the researcher's knowledge, there is no existing study that has dealt with the predictive validity of high school grade point average and college entrance test scores in the Republic of Yemen. Hence, this study serves as a baseline for further research to better understand the nature of the predictive validity of the two predictor variables, high school GPA and college entrance test scores.

Giving the fact that a great amount of variance in the criteria variables is unexplained by the predictor variables in this study, there is room for more research to study the unexplained portion of academic performance. Admission committees can admit students based on other factors. These factors could include a total score based on three years of high school, motivation and interest, orientation, study habits, high school class size, and socioeconomic status to account for more variance.

High school grade point average was not found to be a significant predictor of students' persistence at any of the nine colleges and the addition of college entrance test scores to the prediction equation improved the prediction of college success at only four colleges. Research on predicting persistence from high school GPA and college entrance test scores claimed that persistence in college is influenced substantially by nonacademic factors and that high school GPA and college entrance test scores are not good predictors. Research examining what factors

are related to students' persistence can help inform the admission process so that universities select students who are more likely to stay and succeed at their institutions.

This study examined differential prediction across gender and high school location; further research on the main causes of gender and high school location differences in prediction needs to be conducted to better understand this source of bias. Future research could also examine whether the predictive validity of first-year college GPA using high school GPA and college entrance test scores is consistent across regions (e.g., costal, interior), high school type (e.g., public, private), and high school major (e.g., Arts, Science).

Lastly, college entrance exams at Yemeni universities are designed to measure what students have learned in high schools rather than the necessary skills (e.g., problem solving, reasoning) needed to perform well in college. A study evaluating the content of college entrance exams relative to what skills are required in college is worth doing because succeeding in college requires students to have high levels of content knowledge as well as core academic skills. Measuring students' content knowledge helps to only identify whether applicants have been exposed to content that prepares them for introductory college courses.

APPENDIX A

EXEMPT LETTER FROM THE UNIVERSITY OF PITTSBURGH INSTITUTE OF REVIEW BOARD

University of Pittsburgh
Institutional Review Board

3500 Fifth Avenue
Pittsburgh, PA 15213
(412) 383-1480
(412) 383-1508 (fax)
http://www.irb.pitt.edu

Memorandum

To: Abdulghani Al-Hattami

From: Sue Beers, PhD, Vice Chair

Date: 10/29/2010

IRB#: PRO10100216

Subject: Differential Predictive Validity of High School GPA and College Entrance Test Scores for University Students in Yemen

---

The above-referenced project has been reviewed by the Institutional Review Board. Based on the information provided, this project meets all the necessary criteria for an exemption, and is hereby designated as "exempt" under section
45 CFR 46.101(b)(4).

Please note the following information:

- If any modifications are made to this project, use the " **Send Comments to IRB Staff**" process from the project workspace to request a review to ensure it continues to meet the exempt category.
- Upon completion of your project, be sure to finalize the project by submitting a "**Study Completed**" report from the project workspace.

Please be advised that your research study may be audited periodically by the University of Pittsburgh Research Conduct and Compliance Office.

APPENDIX B

HIGH SCHOOL SUBJECTS FOR THE SCIENCE AND LITERARY TRACKS

Table 29. High school subjects for the Science and Literary tracks

Subject	First secondary	Second secondary		Third secondary	
		Science	Literary	Science	Literary
Holy Quran	✓	✓	✓	✓	✓
Islamic Education	✓	✓	✓	✓	✓
Arabic	✓	✓	✓	✓	✓
English	✓	✓	✓	✓	✓
Mathematics	✓	✓	-	✓	-
Physics	✓	✓	-	✓	-
Chemistry	✓	✓	-	✓	-
Biology	✓	✓	-	✓	-
History	✓	-	✓	-	✓
Geography	✓	-	✓	-	✓
Society	✓	-	-	-	-
Sociology	-	-	✓	-	-
Economics	-	-	✓	-	-
Statistics	-	-	✓	-	✓
Psychology	-	-	-	-	✓
Philosophy	-	-	-	-	✓
Logic	-	-	-	-	✓

APPENDIX C

UNIVERSITIES AND THEIR YEAR OF ESTABLISHMENTS

Table 30. University Names, Number of Colleges and Departments according to the Humanities and Applied Disciplines until the End of 2007/2008

University Name	Establish-ment	Humanities Colleges	Applied Colleges	Number of Colleges	Number of programs
Sana'a	1970	15	11	26	104
Aden	1970	14	6	20	92
Taiz	1993	4	4	8	77
Hodeidah	1996	6	5	11	48
Ibb	1996	4	4	8	35
Dhamar	1996	6	8	12	52
Hadramout	1996	7	6	13	59
Amran	2005	6	1	7	36
Total		60	45	105	503

APPENDIX D

THE NUMBER OF ENROLLED AND GRADUATED STUDENTS IN GENERAL EDUCATION SYSTEM AND PUBLIC UNIVERSITIES FOR THE PERIOD 2002/2003-2007/2008

Table 31. Students Enrollment in the General Education System (Basic and Secondary Education) during the Period 2003/2004-2007/2008

Academic Year	Enrollment				Total
	Male	Percentage	Female	Percentage	
2003/2004	2,836,461	62.4%	1,708,285	37.6%	4,544,746
2004/2005	2,859,303	61.3%	1,805,418	38.7%	4,664,721
2005/2006	2,717,051	60.4%	1,780,592	39.6%	4,497,643
2006/2007	2,882,700	59.4%	1,968,415	41.6%	4,851,115
2007/2008	2,795,875	58.9%	1,954,713	42.1%	4,750,588

Source: Ministry of Education, Yemen 2008 Statistical Year Book.

Table 32. Students Enrollment in General Education by Gender and High School Location during the Periods 2002/2003 and 2007/2008

High Sch. Location	2003/2004		Total	Female %	2007/2008		Total	Female %
	Male	Female			Male	Female		
Rural	787,776	652,688	1,440,464	45.31	804,041	685,418	1,489,459	46.02
%	27.77	38.21	31.70	-	28.76	35.06	31.35	-
Urban	2,048,685	1,055,597	3,104,282	34.00	1,991,834	1,269,295	3,261,129	38.92
%	72.23	61.79	68.30	-	71.24	64.94	68.65	-
Total	2,836,461	1,708,285	4,544,746	37.59	2,795,875	1,954,713	4750588	41.15

Source: Ministry of Education, Yemen 2008 Statistical Year Book.

Table 33. The Number of Enrolled Students in Public Universities for 2002/2003 – 2007/2008

University	2002/2003	2003/2004	2004/2005	2005/2006	2006/2007	2007/2008
Sana'a	84,693	84,816	81,478	81,017	88,615	69,256
Aden	22,547	22,713	23,819	28,367	28,128	28,223
Taiz	29,208	26,061	25,175	23,412	25,729	27,271
Hodeidah	15,317	14,620	12,705	13,069	14,549	15,845
Ibb	9,869	9,366	9,335	9,023	9,563	10,054
Dhamar	13,264	11,290	11,734	12,031	13,713	14,876
Hadramout	6,452	6,519	6,877	7,116	8,260	9,750
Amran	-	-	-	-	-	23,993
Total	181,350	175,385	171,123	174,035	188,557	199,268
Female	46,354	46,455	45,857	48,459	54,601	60,828
%	25.6%	26.5%	26.8%	27.8%	29%	30.5%

Source: Ministry of Higher Education and Scientific Research.

Table 34. The Number of Graduated Students from Public Universities during 2002/2003-2007/2008

University	2002/2003	2003/2004	2004/2005	2005/2006	2006/2007	2007/2008
Sana'a	5,657	9,620	9,115	9,633	9,440	5,396
Aden	1,982	3,848	4,471	5,138	4,305	5,250
Taiz	3,511	3,844	4,548	3,050	3,484	3,484
Hodeidah	1,579	1,781	2,135	1,384	1,907	2,090
Ibb	961	1,314	1,720	1,714	1,202	1,202
Dhamar	1,342	1,839	1,812	1,271	1,356	1,408
Hadramout	1,093	1,083	1,286	1,167	1,100	1,382
Amran	-	-	-	-	-	1,485
Total	16,125	23,329	25,087	23,357	22,794	21,697
Female	5,707	7,832	8,216	7,481	7,884	7,948
%	35.4%	33.6%	32.8%	32%	34.6%	36.6%

Source: Ministry of Higher Education and Scientific Research.

APPENDIX E

COLLEGES AND THEIR YEAR OF ESTABLISHMENT

Table 35. Colleges in Sana'a University and their Year of Establishment (26 colleges)

College	Established	College	Established
Shariah and Law	1970/1971	Arts in Saada	1993/1994
Education in Sana'a	1970/1971	Education and Science in Saada	1993/1994
Arts	1973/1974	Education in Arhab	1993/1994
Science	1973/1974	Education in Amran	1995/1996
Commerce and Economics	1975/1976	Education and Science in Khawlan	1997/1998
Medicine and Health Sciences	1983/1984	Arts in Khawlan	1997/1998
Engineering	1983/1984	Pharmacy	2001/2002
Agriculture	1984/19857	Dentistry	2001/2002
Commerce in Khamer	1995/1996	Education in Abss	2002/2003
Information	1996/1997	Education and Arts and Sciences in Ma'rib	2004/2005
Languages	1997/1998	Languages in Amran	2005/2006
Education in Hajjah	1989/1990	Applied Science in Hajjah	2005/2006
Education in Mahweet	1992/1993	Computer and Information Technology	2007/2008

Table 36. Colleges in University of Aden and their Year of Establishment (20 colleges)

College	Established	College	Established
Education in Aden	1970	Arts	1995
Nasser's Faculty of Agricultural Sciences	1972	Oil and Minerals in Shabwa	1996
Economics	1974	Education in in Lodar	1998
Medicine and Health Sciences	1975	Education in in Dhala	1998
Engineering	1978	Education in Radfan	1998
Law	1978	Education in Yafa'a	1998
Education in Zingibar	1979	Education in Toor Al-Baha	2000
Education in Saber	1980	Management	2000
Education in Shabwa	1994	Administrative Sciences	2000
Dentistry	2009	Pharmacology	2009

Table 37. Colleges in Taiz University and their Year of Establishment (8 colleges)

College	Established
Education in Taiz	1985/1986
Science	1990/1991
Arts	1991/1992
Administrative Sciences	1994/1995
Law	1997/1998
Education, Science and Arts in Al-Turba	1999/2000
Medicine and Health Sciences	1999/2000
Engineering and Information Technology	2003/2004

Table 38. Colleges in Dhamar University and their Year of Establishment (12 colleges)

College	Established	College	Established
Education in Dhamar	1990/1991	Computing and Information Systems	1997/1998
Education and Science in Rda'a	1993/1994	Administrative Sciences	1997/1998
Applied Science	1996/1997	Medicine and Health Sciences	1998/1999
Agriculture and Veterinary Medicine	1996/1997	Engineering	1998/1999
Dentistry	1997/1998	Education in Baida	1998/1999
Arts and languages	1996/1997	Dentistry	2001/2002

Table 39. Colleges in Hodeidah University and their Year of Establishment (11 colleges)

College	Established	College	Established
Education in Hodeidah	1987/1988	Physical Education	1997/1998
Education in Zabid	1992/1993	Fine Arts	1997/1998
Commerce and Economics	1996/1997	Medical Science	1998 / 1999
Shariah and Law	1996/1997	Computer Science and Engineering	2000/2001
Marine and Environmental Sciences	1996/1997	Dentistry	2005/2006
Arts	1996/1997		

Table 40. Colleges in Ibb University and their Year of Establishment (8 colleges)

College	Established
Education in Ibb	1988/1989
Education in Nadera	1993/1994
Arts	1996/1997
Science	1996/1997
Agriculture and Veterinary Medicine	1996/1997
Commerce and Administrative Sciences	1996/1997
Engineering and Architecture	1998/1999
Dentistry	2000/2001

Table 41. Colleges in Hadramout University for Science and Technology and their Year of Establishment (13 colleges)

College	Established	College	Established
Education in Al-Mukalla	1994/1995	Girls College of Education	1997/1998
Engineering and Petroleum	1995/1996	Education in Al-Mahrah	1998/1999
Environmental Science and Marine Biology	1996/1997	Education in Socotra	2000/2001
Medicine and Health Sciences	1997/1998	Arts	2005/2006
Administrative Sciences	1997/1998	Science	2005/2006
Applied Science	1997/1998	Nursing in Saiyon	2005/2006
Education in Saiyon	1996/1997		

APPENDIX F

COLLEGES AND THEIR REQUIRED HIGH SCHOOL GPA AND COLLEGE ENTRANCE TEST SCORES AT HODEIDAH AND IBB UNIVERSITIES

Table 42. Colleges and their Required High School GPA and College Entrance Tests for Hodeidah University

College	HSP	Entrance Tests
Medical Science	85%	English, Chemistry, Biology
Dentistry	85%	English, Physics, Chemistry, Biology
Computer Science and Engineering	80%	English, Physics, Math
Education	75%	English (only in the English Dept.)
Commerce and Economics	75%	N/A
Marine and Environmental Sciences	75%	N/A
Physical Education	75%	N/A
Arts	70%	English (only in the English Dept.)
Fine Arts	70%	N/A
Shariah and Law	70%	N/A

Table 43. Colleges and their Required High School GPA and College Entrance Tests at Ibb University

College	HSP	Entrance Tests
Sciences	75%	Chemistry, Biology and Math
Dentistry	85%	English, Physics, Chemistry, Biology
Engineering and Architecture	80%	English, Physics, Math
Education (Ibb)	75%	English (for English Dept.) and English, Physics, Computer (for Computer Dept.)
Arts	75%	English (only in the English Dept.)
Education (Nadera)	75%	English (for English Dept.) and English, Physics, Computer (for Computer Dept.)
Agriculture and Veterinary Medicine	70%	N/A
Commerce and Administrative Sciences	70%	N/A

APPENDIX G

GRADING SYSTEM AT YEMENI UNIVERSITIES

Table 44. Grading System at Yemeni Universities

Scores	GPA	Grades	Symbol
From 90 to 100	From 4.5 to 5.00	Excellent	A
From 80 to 90	From 4.00 to 4.5	Very good	B
From 65 to 80	From 3.25 to 4.00	Good	C
From 50 to 65	From 2.5 to 3.25	Accepted	D
Less than 50	Less than 2.5	Failure	E

APPENDIX H

FREQUENCY DISTRIBUTION OF THE PARTICIPANTS BY GENDER AND HIGH SCHOOL LOCATION AT HODEIDAH AND IBB UNIVERSITIES

Table 45. Frequency Distribution of the Participants by Gender at Hodeidah University

College	N	Gender	Frequency	Percentage
Arts	45	Female	30	66.67
		Male	15	33.33
Computer Science and Engineering	96	Female	31	32.29
		Male	65	67.71
Dentistry	67	Female	50	74.63
		Male	17	25.37
Education	107	Female	79	73.83
		Male	28	26.17
Medical Science	187	Female	91	48.66
		Male	96	51.34

Table 46. Frequency Distribution of the Participants by Gender at Ibb University

College	N	Gender	Frequency	Percentage
Arts	166	Female	80	48.19
		Male	86	51.81
Dentistry	68	Female	42	61.76
		Male	26	38.24
Engineering and Architecture	51	Female	20	39.22
		Male	31	60.78
Sciences	94	Female	38	40.43
		Male	56	59.57

Table 47. Frequency Distribution of the Participants by High School Location at Hodeidah University

College	N	Gender	Frequency	Percentage
Arts	45	Rural	21	46.67
		Urban	24	53.33
Computer Science and Engineering	96	Rural	38	39.58
		Urban	58	60.42
Dentistry	67	Rural	11	16.42
		Urban	56	83.58
Education	107	Rural	44	41.12
		Urban	63	58.88
Medical Science	187	Rural	63	33.69
		Urban	124	66.31

Table 48. Frequency Distribution of the Participants by High School Location at Ibb University

College	N	Gender	Frequency	Percentage
Arts	166	Rural	66	39.76
		Urban	100	60.24
Dentistry	68	Rural	25	36.76
		Urban	43	63.24
Engineering and Architecture	51	Rural	23	45.10
		Urban	28	54.90
Sciences	94	Rural	30	31.91
		Urban	64	68.09

APPENDIX I

PREDICTORS AND CRITERION VERIABLES INTERCORRELATIONS FOR EACH COLLEGE AND BY GENDER AND HIGH SCHOOL LOCATION

Table 49. Intercorrelations among the variables across all colleges

	HSL	HSGPA	CETS	FYGPA	CGPA	Grad.
Gender	-0.040*	-0.068**	-0.140**	-0.093**	-0.062**	0.052**
HSL		0.021	0.029	0.020	0.005	0.116**
HSGPA			0.272**	0.409**	0.372**	-0.054**
CETS				0.417**	0.410**	0.010
FYGPA					0.698**	-0.013
CGPA						na

HSL = High School Location; HSGPA = High School Grade Point Average; CETS = College Entrance Test Scores; FYGPA = First Year Grade Point Average; CGPA = Cumulative Grade Point Average; Grad. = Graduated.
* = $P < 0.05$; ** = $P < 0.01$

na = not applicable because graduated variable is constant. All those who obtained cumulative GPA were graduated.

Table 50. Intercorrelations among the variables across all by gender (Upper Triangle for Female)

	HSGPA	CETS	FYGPA	CGPA	Grad.
HSGPA		0.276**	0.435**	0.404**	-0.049
CETS	0.273**		0.471**	0.478**	0.022
FYGPA	0.380**	0.353**		0.750**	-0.026
CGPA	0.342**	0.341**	0.653**		na
Grad.	-0.054*	0.008	0.009	na	

HSL = High School Location; HSGPA = High School Grade Point Average; CETS = College Entrance Test Scores; FYGPA = First Year Grade Point Average; CGPA = Cumulative Grade Point Average; Grad. = Graduated.
* = P < 0.05; ** = P < 0.01

na = not applicable because graduated variable is constant. All those who have cumulative GPA were graduated.

Table 51. Intercorrelations among the variables for all by high school location (Upper Triangle for Rural)

	HSGPA	CETS	FYGPA	CGPA	Grad.
HSGPA		0.233**	0.406**	0.397**	-0.023
CETS	0.291**		0.358**	0.361**	0.007
FYGPA	0.410**	0.448**		0.704**	0.024
CGPA	0.358**	0.435**	0.695		na
Grad.	-0.088**	0.009	-0.049	na	

HSL = High School Location; HSGPA = High School Grade Point Average; CETS = College Entrance Test Scores; FYGPA = First Year Grade Point Average; CGPA = Cumulative Grade Point Average; Grad. = Graduated.
* = P < 0.05; ** = P < 0.01

na = not applicable because graduated variable is constant. All those who have cumulative GPA were graduated.

BIBLIOGRAPHY

Abbass, A. (2011). *The Prevailing Regulatory System in High Schools and Its Impact on Students Achievement in Yemen*. Ph.D. Dissertation. Algeria University. http://www.manaralyemen.com/news-8190.html.

ACT, Inc. (1997). *ACT Assessment Technical Manual*. Iowa City, IA: Author.

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment*. Washington, DC: U.S. Dept. of Education.

Aiken, L. R. (1997). *Psychological testing and assessment (9th ed.)*. Boston: Allyn & Bacon.

Alfred, R. (1973). *Student Attrition strategies for action. The proceedings of a conference sponsored by GT-70.* Kansas City, MO: Office of Educational Planning and Research. Metropolitan Junior College District, Kansas City Office.

Alkhateeb, A. (1992). Higher Education At the Arab World: Challenges and future alternatives. *Journal of the Union of Arab Universities, 27*, 91-120.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1995). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

Anastasi, A. (1988). *Psychological Testing (6th ed.)*. New York: Macmillan.

Anastasi, A., & Urbina S. (1997). *Psychological Testing (7th ed.)*. New York: McMillian.

Armstrong, W.B. (1999). The relationship between placement testing and curricular content in the community college: Correspondence or misalignment? *Journal of Applied Research in the Community College, 7*(1), 33-38. Retrieved August 30, 2010, from http://www.rpgroup.org/publications/ConfProceedings-WorkshopPapers/38th-conf-may-2000/02-Armstrong-PlacementTesting-8B-Curricular-Content.Pdf.

Astin, A. W. (1977). *Four critical years*. San Francisco: Jossey-Bass.

Astin, A. W. (1997). How "good" is your institution's retention rate? *Research in Higher Education, 38*, 647-658.

Astin, A. W., Korn W., & Green, K. (1987). Retaining and satisfying students. *Educational Record, 68*(1), 36-42.

Azen, R., Bronner, S., & Gafni, N. (2002). Examination of gender bias in university admissions. *Applied Measurement in Education, 15*, 75-94.

Bachman, L. F., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Baird, L. L. (1984). Predicting predictability: The influence of student and institutional characteristics on the prediction of grades. *Research in Higher Education, 21*, 261-279.

Barnett, J., Ritter, G. W., & Lucas, C. J. (2003). Educational reform in Arkansas: Making sense of the debate over school consolidation. *Public Policy Working Paper Series, 2*(1), 1-27.

Baron, J., & Norman, F. (1992). SATs, achievement tests, and high school class rank as predictors of college performance. *Educational and Psychological Measurement, 52*, 1047-1055.

Beecher, M., & Fischer, L. (1999). High school courses and scores as predictors of college success. *The Journal of College Admission*, *163*, 49.

Berdie, R. F. (1960). Some principles and problems of selective college admissions: Should entrance requirements be raised. *Journal of Higher Education, 31*(4), 191-199.

Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). Equal Resources, Equal Outcomes? The Distribution of School Resources and Student Achievement in California. San Francisco: Public Policy Institute of California.

Bohrnstedt, G. W., & Stecher, B. M. (Eds). (1999). *Class size reduction in California: Early evaluation findings*, 1996-1998. (CSR Research Consortium, Year 1 Evaluation Report). Palo Alto, CA: American Institutes for Research.

Bontekoe, J. F. (1992). *The ACT as a predictor of college success at Trinity Christian College* (Report # TM 079599) Palos Heights, II: Trinity Christian College.

Bowen, W. & Bok, D. (1998). *The shape of the river: Long term consequences of considering race in college and university admissions.* Princeton: Princeton UP.

Braxton, J. M. (2003) Persistence as an essential gateway to student success. In, S. Komives & D. Woodard (Eds), *Student services: a handbook for the* profession *(4th ed)*, 317-335. San Francisco: Jossey-Bass.

Breland, H. M. (1979). *Population Validity and College Entrance Measures* (Research Monograph No. 8). New York: College Entrance Examination Board.

Breland, H., Maxey, J., Gernand, R., Cumming, T. & Trapani, C. (2002) *Trends in college admission 2000: A report of a national survey of undergraduate admission policies, practices, and procedures.* Retrieved January 20, .2010, from http://www.airweb.org/images/trendsreport.pdf.

Bridgeman, B. & McHale, F. (1996). *Gender and ethnic group differences on the GMAT Analytical Writing Assessment* (ETS RR-96-2). Princeton, NJ: Educational Testing Service.

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Predictions of freshman grade point average from the revised and recentered SAT I: Reasoning Test* (College Board Rep. No. 2000-1). New York: College Entrance Examination Board.

Burton, N. W., & Ramist, L. (2001). *Predicting long term success in undergraduate school: A review of predictive validity studies*. Princeton, NJ: Educational Testing Service.

Bylund, R. A., & Reeves, E. B. (2005). Are Rural Schools Inferior to Urban Schools? A Multilevel Analysis of School Accountability Trends in Kentucky. *Rural Sociology, 70*(3), 360-384.

Cabrera, A. F., Nora, A., & Castañeda, M. B. (1993). College persistence: Structural equations modeling test of an integrated model of student retention. *Journal of Higher Education, 64*, 123–139.

Camara, W. J., & Echternacht, G. (2000). The SAT I and high school grades: Utility in predicting success in college. *College Board Research Notes* (RN-10). New York: College Board.

Camara, W., Kimmel, E., Scheuneman, J., & Sawtell, E.A. (2003). Whose Grades Are Inflated? *College Board Research Report No. 2003-4*. Retrieved from http://www.collegeboard.com/research/pdf/04843cbreport20034_31757.pdf.

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: SAGE Publications.

Carmines, E. G. & Zeller, R. A. (1991). *Reliability and validity assessment*. Newbury Park: Sage Publications.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.

Cleary, T. A., Humphreys, L. G., Kendrick, S. A. & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist, 30*, 15-41.

Cole, N. S. (1972). *Bias in selection* (ACT Research Rep. No. 51). Iowa City, IA: The American College Testing Program.

Constitution of the Republic of Yemen. (1994). *Government of the Republic of Yemen*. Retrieved February 4, 2009, from: http://www.al-bab.com/yemen/gov/con94.htm.

Cowen, S., & Fiori, S. J. (1991, November). *Appropriateness of the SAT in selecting students for admission to California State University, Hayward*. Paper presented at the annual meeting of the California Educational Research Association, San Diego, CA.

Crede', M., & Kuncel, N. R. (2008). Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance. *Perspectives on Psychological Science*, *3*, 425-453.

Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modem Test Theory*. New York: Holt, Rinehart and Winston.

Cronbach, L. (1971). Test validation. In R. Thorndike (Ed.), *Educational measurement (2nd ed.)*, 443-507. Washington, DC: American Council on Education.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Cronbach, L. J., Schoeneman, P., & McKie, D. (1965). Alpha coefficients for stratified tests. *Educational and Psychological Measurement, 25*, 291-312.

Crouse, J., & Trusheim, D. (1988). *The case against the SAT*. Chicago: University of Chicago.

Dresch, P. (2000). *A History of Modern Yemen.* Cambridge: Cambridge University Press.

Duran, R. P. (1986). Prediction of Hispanics' college achievement. In M. A. Olivas (Ed.), *Latino college students*, 221-245. New York: Teachers College Press.

Ebel, R. L. (1978, October). *The role of standardized tests in college admissions, credit and placement*. Paper presented at the National Forum of College Entrance Examination Admission, New York.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement (5th ed.)* Englewood Cliffs, NJ: Prentice Hall.

Education Indicators in the Republic of Yemen, National Information Center. Retrieved February 4, 2011, from: http://www.yemennic.info/contents/education/bookses/index.php.

Educational Testing Service (1999). *Overview: ETS fairness review*. Author. Retrieved August 30, 1999 from http://www.ets.org.

Eimers, M. T. & Pike, G. R. (1997). Minority and nonminority adjustment to college: Differences or similarities? *Research in Higher Education, 38*(1), 77-97.

Elert, G. (1992). The SAT, aptitude or demographics? E-World. Retrieved March 17, 2011, from: http://www.hypertextbook.com/eworld/.

Elliott, R., & Strentra, A. C. (1988). Effects of improving the reliability of the GPA on prediction generally and on comparative predictions for gender and race particularly. *Journal of Educational Measurement, 25*, 333-347.

Encyclopedia Britannica. Retrieved February 4, 2009, from: http://www.britannica.com/EBchecked/topic/652831/Yemen.

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43*, 38290-38315.

Feldt, L. S., & Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement (3rd ed.)*, 105-146. Phoenix, AZ: Oryx Press.

Fleming, J. (2002). Who will succeed in college? When the SAT predicts Black students' performance. *Review of Higher Education, 25*(3), 281-296.

Fleming, J., & Garcia, N. (1998). Are standardized tests fair to African Americans? Predictive validity of the SAT in Black and White institutions. *Journal of Higher Education, 69*(5), 471-495.

Fontaine, R. & Exum, A. (January 5, 2010). Yemen's coming disaster. Its oil is expected to run out in 2017, but Yemen hasn't planned for its young, poverty-ridden population's post-oil future. *Los Angeles Times*.

Friedman, B. A., and Mandel, R. G.(2010). The Prediction of College Student Academic Performance and Retention: Application of Expectancy and Goal Setting Theories. *Journal of college student retention: Research, theory & practice, 11*(2), 227-246.

General Law for Education, 1992. Retrieved February 4, 2010, from: http://www.moeyemen.net/About/DecisionsAndLaws/Laws.aspx.

Geske, A., Grinfelds, A., Dedze, I., & Zhang, Y. (2006). Family background, school quality and rural-urban disparities in student learning achievement in Latvia. *Prospects. 36*(4), 419-431.

Giordano, F. G. (1996). Evaluation as empowerment: Using evaluation strategies to improve retention of regularly admitted students of color. *New Directions for Student Services, 74*, 69-77.

Gose, B. (1994). Study Says Rigor of High School Coursework is the Best Predictor of College Graduation. *The Chronicle of Higher Education, XLV* (39), A46.

Grissmer, D., Flanagan, A., Kawatka, J., & Williamson, S. (2000). *Improving student achievement: What state NAEP scores tell us.* Santa Monica: Rand.

Gulliksen, H., & Wilks, S. S. (1950). Regression tests for several samples. *Psychometrika, 15*, 91-114.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational Measurement (4th ed.)*, 65-110. Westport, CT: American Council on Education and Praeger.

Helmstadter, G. C. (1964). *Principles of psychology measurement*. New York: Appleton Century- Crofts.

Hezlett, S. A., Kuncel, N., Vey, M. A., Ahart, A. M., Ones, D. S., Campbell, J. P., & Camara, W. J. (2001, April). *The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis*. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, WA.

Higher Education National Strategy, 2009. Retrieved February 4, 2010, from: http://www.yemen-nic.info/contents/education/bookses/pdf/3-.pdf.

Higher Education Project II, 2008. Retrieved February 4, 2010, from: http://www.yemen-nic.info/contents/education/bookses/pdf/10-.pdf.

Hill, G. (2007). *Somalis pin peace hopes on Yemen*. BBC News. April 17, 2010.

Himelstein, P. (1965). Validities and Inter-correlations of MMPI Subscales Predictive of College Achievement. *Educational and Psychological Measurement, 4*, 1125-1128.

Hirschberg, N. W. (1977). Predicting performance in graduate school. In M.F. Kaplan & S. Schwartz (Eds.).*Human judgment and decision processes in applied settings,* 95-123. New York: Academic Press.

Hoffman, J. L. (2002). The impact of student co-curricular involvement on student success: Racial and religious differences. *Journal of College Student Development, 43*(5), 712-739.

Hoffman, J. L., & Lowitzki, K. E. (2005). Predicting college success with high school grades and test scores: Limitations for minority students. *The Review of Higher Education, 28 (4),* 455-474.

House, J. D., & Keeley, E. J. (1993, November). *Differential prediction of graduate student achievement from Miller Analogies Test scores*. Paper presented at the annual meeting of the Illinois Association for Institutional Research, Oakbrook Terrace, IL.

Hu, N. B. (2002). *Measuring the Weight of High school GPA and SAT Scores with Second Term GPA to Determine Admission/ Financial Aid*. Paper Presented at the Annual Meeting of the Association for Institutional Research. 42nd, Toronto, Ontario, Canada.

Human Development Report 2005, UNDP, New York. Retrieved February 4, 2009, from: http://hdr.undp.org/en.

Humphreys, L. G. (1962). The nature and organization of human abilities. *The Yearbook of the National Council on Measurement in Education, 19,* 16-19.

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*, 53-69.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender difference in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Ibrahim, R., A., (2006). Making Aptitude Test for High School Students in the Republic of Yemen. National Center for information.

Ishitani, T. T. (2005). *Studying Educational Attainment among First-Generation Students in the United States.* Paper presented at the 45th Annual Forum of the Association for Institutional Research, San Diego, CA.

Ishitani, T. T. (2006). Studying attrition and degree completion behavior among first generation college students in the United States. *The Journal of Higher Education, 77*(5), 861-885.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Johnson, J. D. (2002). *Size, excellence, and equity: A report on Arkansas schools and districts*. Athens, OH: Ohio University College of Education.

Johnson, J. W., Carter, G. W., Davison, H. K. , & Oliver, D. H. (2001). A synthetic validity approach to testing differential prediction hypotheses. *Journal of Applied Psychology, 86*, 774-780.

Kamata, A., Turhan, A., & Darandari, E. (2003). *Estimating Scale Reliability of Multidimensional Composite Scores*. Paper presented at the annual meeting of American Educational Research Association, Chicago, April 2003.

Kerlinger, F. N. & Pedhazur, E. J. (1973). *Multiple regression in behavioral research*. New York: Holt, Rinehart, and Winston.

Kim, M. M. (2002). Historically Black vs. White institutions: Academic development among Black students. *Review of Higher Education, 25*(4), 385-407.

King, B. W., Rasool, J. A., & Judge, J. J. (1994). The relationship between college performance and basic skills assessment using SAT scores, the Nelson Denny Reading Test and degrees of reading power. *Research and Teaching in Developmental Education, 11*(1), 5-13.

Kleinfeld, J. S., et al. (1985). *Alaska's small rural high schools: are they working?* ISER Report Series No. 58. Fairbanks: University of Alaska.

klugh, H. E., & Bierly, R. (1959). The School and College Ability Test and High School Grades as Predictors of College Achievement. *Educational and Psychological Measurement, 4*, 625-626.

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science, 315*, 1080-81.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2007). A comprehensive meta-analysis of the predictive validity of the Graduate Management Admission Test (GMAT) and undergraduate grade point average (UGPA). *Academy of Management Learning and Education, 6*, 51-68.

Kuncel, N. R., Credé, M., Thomas, L. L., Seiler, S. N., Klieger, D. M., & Woo, S. E. (2005).A meta-analysis of the Pharmacy College Admission Test (PCAT) and grade predictors of pharmacy student success. *American Journal of Pharmaceutical Education, 69*, 339-347.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin, 127,* 162-181.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86,* 148-161.

Lawlor, S., Richman, S. & Richman, C. L. (1997). The Validity of Using the SAT as a Criterion for Black and White Students' Admission to College. *College Student Journal. 31*, 507-515.

Lenning, O. T. (1975). *Predictive validity of the A CT tests at selective colleges*. Iowa City, Iowa: The American College Testing Program.

Levin, J., & Wyckoff, J. H. (1994). Predicting persistence and success in baccalaureate engineering. *Education, 111*, 461-468.

Lin, Tin-Chun; Yu, Wei-Choun; Chen, Yi-Chi; and Kauffman, Daniel, (2011). Determinants and Probability Prediction of College Student Retention: New Evidence from the Probit Model. *International Journal of Education Economics and Development (indexed by EconLit*), Under Review.

Linn, R. L. (1973). Fair test use in selection. *Review of Educational Research, 43*, 139-161.

Linn, R. L. (1978). Single-group validity, differential validity, and differential prediction. *Journal of Applied Psychology, 63,* 507–512.

Linn, R. L. (1982). Ability testing: Individual differences, prediction, and differential prediction. In A. Wigor & W. Garner (Eds.), *Ability testing: Uses, consequences, and controversies* (pp. 335-338). Washington, DC: National Academy Press.

Linn, R. L. (1982a). Admissions testing on trial. *American Psychologist, 37,* 279–291.

Linn, R. L. (1982b). Ability testing: Individual differences, prediction and differential prediction. In R. L. Linn (Ed.), *Ability testing: Uses, consequences, and controversies.* Washington, DC: National Academy Press.

Linn, R. L. (1984). Selection bias: Multiple meanings. *Journal of Educational Measurement, 21*, 33-48.

Linn, R. L. (1990). Admissions testing: Recommended uses, validity, differential prediction, and coaching. *Applied Measurement in Education, 34*, 297-318.

Linn, R. L. (1993). The use of differential item functioning statistics: A discussion of current practices and future implications. In P. Holland & H. Wainer (Eds.), *Differential item functioning*, 349-364. Hillsdale, NJ: Lawrence Erlbaum.

Linn, R. L. (1994). Fair test use: Research and policy. In M. G. Rumsey, C. B. Walker, & J. H. Harris (Eds.), *Personnel Selection and Classification*, 363–375. Hillsdale, NJ: Lawrence Erlbaum.

Littleton, R. A. (2001). *African American student persistence at small colleges in Appalachia: A qualitative study*. Paper Presented at Association for the Study of Higher Education (ASHE). Annual Conference in Richmond, Virginia. November 15-18, 2001.

Lohfink, M. M. & Paulsen, M. B. (2005). Comparing the determinants of persistence for first-generation and continuing-generation students. *Journal of College Student Development, 46*(4), 409-428.

Maccoby, E. E. & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University.

Mathiasen, R. E. (1984). Predicting college academic achievement: a research review. *College Student Journal, 18*, 380-386.

McDonald, R. P. (1999). *Test theory: A unified treatment.* Hillsdale, NJ: Erlbaum.

Mehrens, W. A., & Lehmann, I. J. (1987). *Using standardized tests in education* (4th ed.). New York: Longman.

Mertler, C. A., & Vannatta, R. A. (2002). *Advanced and Multivariate Statistical Methods* (2nd ed.). Los Angeles, CA: Pyrczak Publishing.

Messick, S. (1989). Validity. In R. L. Linn, (ed.), *Educational Measurement*. New York: Macmillan Publishing Company, 13-103.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50(9),* 741-749.

Messick, S. (2000). Consequences of test interpretation and use: The fusion of validity and values in psychological assessment. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment* (3-20). Norwell, MA: Kluwer Academic.

Micceri, T. (2001). Facts and fantasies regarding admission standards (Report No. HE034083). Long Beach, CA: Association for Institutional Research.

Ministry of Education, Illiteracy and Adult Education Strategy, 2008. Retrieved February 4, 2010, from: http://www.yemen-nic.info/contents/education/bookses/index.php.

Ministry of Education, Yemen 2008 Statistical Year Book. Retrieved February 4, 2010, from: http://www.cso-yemen.org/.

Ministry of Education. (2010). Retrieved January 12, 2010, from: www.moeyemen.net.

Ministry of Education. Statistical Year Book, 2006/2007. Retrieved February 4, 2010, from: http://www.yemen-nic.info/contents/education/bookses/pdf/5-.pdf.

Ministry of High Education and Scientific Research. (2010). Retrieved January 12, 2010, from: www.moheyemen.org.

Misanchuk, E. L. (1977). A model-based prediction of Scholastic Achievement. *Journal of Educational Research, 71*, 30-35.

Moffat, G. K. (1993). *The validity of the SAT as a predictor of grade point average for nontraditional college students*. Paper presented at the annual meeting of the Eastern Educational Research Association, Clearwater Beach, FL. ERIC Document Reproduction Service No. ED 356 252.

Monk, D. H., & Haller, E. J. (1986). *Organizational alternatives for small rural schools.* Cornell: New York State College of Agriculture and Life Sciences at Cornell University, 1986.

Moores, J. (2003). *A preliminary report on the University of California, Berkeley admission process for 2002*. Oakland, CA: UC Office of the President: Downloaded from http://www.universityofcalifornia.edu/news/compreview/mooresreport.pdf.

Morgan, R. (1989). *Analysis of the Predictive Validity of the SAT and High School Grades from 1976 to 1985*. College Board Research Report No. 89-7. New York: The College Board.

Morgan, R. (1990). Analysis of predictive validity within student categorizations. In W. Willingham, C. Lewis, R. Morgan, & L. Raimus (Eds.), *Predicting college grades*, 225 – 238. Princeton, NJ: Educational Testing Service.

Mouw, J. T., & Khanna R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal 27*(4 or 3), 328-336.

Mulvenon, S. W., Stegman, C. E., Ganley, B. J., & McKenzie, S. (2002) *Little Rock School District revisited: Is desegregation working?* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Mulvenon, S. W., Stegman, C., Thorn, A., & Thomas, S. (1999, winter). Selection for college admission: Refining traditional models. *Journal of College Admission, 162*, 20-27.

Munro, B. H. (1981). Dropouts from higher education: Path analysis of a national sample. *American Educational Research Journal, 18*(2), 133-141.

Myers, R. S., & Pyles, M. R. (1992). *Relationships among high school grades, ACT test scores, and college grades*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Knoxville, TN. ERIC Document Reproduction Service No. ED 353 317. National Council of Teachers of Mathematics.

National Information System, 2010. Retrieved February 4, 2010, from: http://www.yemen-nic.info/index.php.

Nettles, A. L., & Nettles, M. T. (Eds.) (1999). *Measuring up: Challenges minorities face in educational assessment*. Norwell, MA: Kluwer Academic.

Nettles, M. (1991). Racial similarities and differences in the predictors of college student achievement. In W. R. Allen & E. G. Epps (Eds.), *College in Black and White: African American students in predominantly White and in historically Black public universities* (pp. 75-94). Albany: State University of New York Press.

Nisbet, J., Ruble, V. E., & Schurr, K. T. (1982). Predictors of academic success with high risk college students. *Journal of College Student Personnel, 57,* 289-318.

Noble, J. P. (1991). *Predicting college grades from ACT assessment scores and high school course work and grade information* (Rep. No. ACT-RR-91-3). Iowa City, IA: American College Testing Program.

Noble, J. P., & Sawyer, R. (1997). *Predicting grades in specific college freshman courses from ACT test scores and self-reported high school grades*. Iowa City, LA.: ACT Research Report Series, 87-20. Retrieved November 3, 2010, from http://www.act.org/research/reports/pdf/ACT_RR87-20.pdf.

Nunnally, J. S. (1978). *Psychometric theory (4th ed.)*. New York: McGraw-Hill.

Odell, C. W. (1927). Attempt at predicting success in freshman year at college. *School and Society, 25*, 702-706.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*(3), 343-355.

Pascarella, E. & Terenzini, P. (1977) Patterns of Student-Faculty Informal Interaction beyond the Classroom and Voluntary Freshman Attrition. *Journal of Higher Education. 48*(5), 540 - 552.

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *Journal of Higher Education, 51*(1), 60-75.

Pascarella, E. T., & Terenzini, P. T. (1991). *How College Affects Students: Findings and Insights from Twenty Years of Research*. San Francisco: Jossey-Bass.

Peltier, G. L., Laden, R., & Matranga, M. (1999). Student Persistence in College: A review of Research. *Journal of Student Retention. 1*(4), 357-375.

Pentages, T. J., & Creedon, C. F. (1978). Studies of college attrition: 1950-1975. *Review of Educational Research, 48*, 49-101.

Perfetto, G. (2002). Predicting academic success in the admissions process: Placing an empirical approach in a larger process. *College Board Review, 196*, 30-35.

Peterson, C. H., Casillas, A., & Robbins, S. B. (2006). The student readiness inventory and the big five: Examining social desirability and college academic performance. *Personality and Individual Differences, 41,* 663-673.

Pike, G. R., & Saupe, J. L. (2002). Does high school matter? An analysis of three methods of predicting first-year grades. *Research in Higher Education, 43*(2), 187-207.

Ramist, L. (1981). *College attrition and retention* (College Board Rep. No. 81-1). New York: College Entrance Examination Board.

Ramist, L. (1984). Predictive validity of ATP tests. In T.F. Donlon (Ed.), *College Board technical handbook for the Scholastic Aptitude and Achievement Tests*, 141-170. New York: College Entrance Examination Board.

Ramist, L., Lewis, C., & McCamley, L. (1990). Implications of using freshman GPA as the criterion for the predictive validity of the SAT. In W. W. Willingham, C. Lewis, R. Morgan& L. Ramist (Eds.), *Predicting college grades: An analysis of institutional trends over two decades,* 253-288. Princeton, NJ: Educational Testing Service.

Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1994). *Student group differences in predicting college grades: Sex, language, and ethnic groups*. College Board Report No. 93-1. New York: College Entrance Examination Board.

Republic of Yemen: Interim Poverty Reduction Strategy Paper, December 2000. Retrieved February 4, 2010, from: http://www.imf.org/external/NP/prsp/2000/yem/01/123100.pdf.

Riehl, R. J. (1994). The academic preparation, aspirations, and first-year performance of first-generation students. *College and University, 70,* 14-19.

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychological and study skill factors predict college outcome? A meta-analysis. *Psychological Bulletin, 130*, 261–288.

Rosenbaum, J. E. (2004, Spring). It's time to tell the kids: If you don't do well in high school, you won't do well in college (or on the job). *American Educator*. Retrieved August 25, 2010, from http://www.aft.org/pubs-reports/american_educator/spring2004/tellthekids.html.

Rowan, R. W. (1978). The predictive value of the ACT at Murray State University over four-year college program. *Measurement and Evaluation in Guidance, 11*(3), 143-149.

Roza, M., Miles, K. H., & Foley, E. (2003). *Within-district spending: A tool for unmasking funding inequities across schools*. Paper presented at the annual meeting of the American Education Finance Association, Orlando, FL.

Rudner, L. (1994). Questions to ask when evaluating tests. *Practical Assessment, Research & Evaluation, 4*(2). Available online: http://PAREonline.net/getvn.asp?v=4&n=2.

Sa'ad, Abdul-Jabbar A. "Educational Challenges Facing Yemen." *Yemen Times VIII 19* (11-17 May 1998).

Sackett, P. R., & Wilk, S. L. (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist. 49 (11)*, 929-954.

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin, 135*, 1-22.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist, 56,* 302-318.

Salvia, J., & Ysseldyke, J. (1995). *Assessment (6th ed.)*. Boston: Houghton Mifflin Company.

Salvia, J., & Ysseldyke, J. (2007). *Assessment in special and inclusive education (10th ed.)*. Boston: Houghton-Mifflin.

Sawyer, R., & Maxey, E. J. (1979). *The validity over time of college freshmen grade prediction equations* (ACT Research Report Series No. 80). Iowa City, IA: ACT, Inc.

Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T., Quinn, A., Sinha, R., & Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact of demographic status on admitted students. *Journal of Applied Psychology, 94,* 1479-1497.

Sedlacek, W. E. (2004). *Beyond the big test: Noncognitive assessment in higher education*. San Francisco: Jossey-Bass.

Seidman, A. (2005). *College Student Retention. Formula for Student Success*. Westport, CT: ACE/Praeger Publishers.

Shea, J. A, & Fortna, G. S. (2002). Psychometric methods. In G. R. Norman, C. P. M. van der Vleuten, & D. I. Newble (Eds.), *International handbook of research on medical education: Part one,* 97- 126. Norwell, MA: Kluwer Academic.

Shepard, L. A. (1982). Definitions of bias. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias* (pp. 9–30). Baltimore, MD: Johns Hopkins University Press.

Shepard, L. A. (1987). Bias in tests of achievement and scholastic aptitude. In S. Modgil & C. Modgil (Eds.), *Arthur Jensen: Consensus and controversy,* 177-190. Philadelphia: Falmer.

Snyder, V., Hackett, R. K., Stewart, M.& Smith, D. (2003). Predicting academic performance and retention of private university freshmen in need of developmental education. *Research & Teaching in Developmental Education, 19*(2), 17-28.

Stanley, J. C., & Porter, A. C. (1967). Correlation of scholastic aptitude scores with college grades for Negros versus Whites. *Journal of Educational Measurement, 4*, 199-218.

Stanley, L. R., Comello, M. L. G., Edwards, R. W., & Marquart, B. S. (2008). School adjustment in rural and urban communities: Do students from "Timbuktu" differ from their "city slicker" peers? *Journal of Youth and Adolescence, 37*(2), 225.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Summers, M. D. (2003). Attrition research at community colleges. *Community College Review, 30*(4), 64-84.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics (4th ed.).* Needham Heights, MA: Allyn & Bacon.

Terenzini, P. T., Springer, L., Yaeger, P. M, Pascarella, E. T., & Nora, A. (1996). First-generation college students: Characteristics, experiences, and cognitive development. *Research in Higher Education, 37*, 1-22.

The College Board (1998, February). *SAT and gender differences* (College Board Research Summary RS-04). New York: The College Board, Office of Research and Development.

The General Census of Population 2004. Yemen News Agency. Retrieved February 4, 2010, from: http://www.sabanews.net/en/news85385.htm. Retrieved 2010-03-21.

Thomas, C. L. (1972, April). *The relative effectiveness of high school grades and standardized test scores for predicting college grades for black students*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago.

Ting, Siu-Man. (1998). First-year grades and academic progress of college students of first-generation and low-income families. *The Journal of College Admission*, Winter, 15-23.

Tinto, V. (1975) Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research, 45*(1), 89-125.

Tinto, V. (1982) Limits of Theory and Practice in Student Attrition. *Journal of Higher Education. 53*(6), 687-700.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago: University of Chicago.

Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of student leaving. *The Journal of Higher Education, 59*(4), 438-455.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. Chicago: The University of Chicago Press.

Tracey, T. J., & Sedlacek, W. E. (1984). Noncognitive variables in predicting academic success by race. *Measurement and Evaluation in Guidance, 16*, 172-178.

Travers, R. M. W. (1994). Significant research on the prediction of academic success. In W. T. Donahue, C. H. Coombs, & R. M. W. Travers (Eds.), The measurement of student adjustment and achievement. Ann Arbor: University of Michigan Press, 147-190.

Tross, S. A., Harper, J. P., Osher, L. W., & Kneidinger, L. M. (2000). Not just the usual cast of characteristics: Using personality to predict college student performance and retention. *Journal of College Student Development, 41*(3), 323-334.

UNESCO, March 2006, The Impact of Women teachers over girl's education: Advocacy brief, Bangkok

Van T. Bui, K. (2002). First-generation college students at a four-year university: Background characteristics, reasons for pursuing higher education, and first-year experiences. *College Student Journal, 36*(1), 3-12. Retrieved November 11, 2010, from http://www.encyclopedia.com/doc/lGl-85007762.html.

Veltri, S., Banning, J., & Davies, T. (2006). The community college classroom environment: Student perceptions. *Community College Journal 40*(3), 517–527.

Wainer, H., & Braun, H. I. (1988). *Test validity*. Hillsdale, NJ: Lawrence Erlbaum.

Warburton, E., Bugarin, R., & Nunez, A. (2001). *Bridging the gap: Academic preparation and postsecondary success of first-generation students*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics, U.S. Government Printing Office. Retrieved November 3, 2010 from http://nces.ed.gov/pubs2001/2001153.pdf.

Waugh, G. & Micceri, T. (1994, March). *Using ethnicity, SAT/ACT scores, and high school GPA to predict retention and graduation rates*. Paper Presented at the Florida Association for Institutional Research Conference. Retrieved April 10, 2010, from http://usfweb2.usf.edu/IR/RePDF/199403_2.pdf.

Waugh, G., Micceri, T., & Takalkar, P. (1994). *Using ethnicity, SAT/ACT scores, and high school GPA to predict retention and graduation rates* (Report No. HE034024). Orlando, FL: Florida Association for Institutional Research. (ERIC Document Reproduction Service No. ED453704).

Whittle, V. E & Denaux, Z. S. (2007) Closing the Rural and Urban Educational Achievement Gap: A Study of the State of Georgia, *Review of Business Research, 7*(5), 60-66.

Wigdor, A. K., & Garner, W. R. (Eds.) (1982). *Ability testing: Uses, consequences, and controversies*. Washington, DC: National Academic.

Wightman, L. F. (2003). Standardized testing and equal access: A tutorial. In M. J. Chang (Ed.), *Compelling interest: Examining the evidence of racial dynamics in colleges and universities*, 49-96. Stanford, CA: Stanford University.

Wilder, G. Z., & Powell, K. (1989). *Sex differences in test performance: A survey of the literature* (College Board Report No. 89-3). Princeton, NJ: Educational Testing Service.

Willingham, W. W. (1974, January 25). Predicting success in graduate education. *Science, 183*, 273-278.

Willingham, W. W. (1985). *Success in college: The role of personal qualities and academic ability*. New York: College Entrance Examination Board.

Willingham, W. W. (1998). Validity in college selection: Context and evidence. Paper presented at the National Academy of Sciences Workshop on the Role of Tests in Higher Education Admissions, Washington, DC.

Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Mahwah: NJ: Erlbaum. (ED 416 293).

Willingham, W., Lewis, C., Morgan, R., & Ramist, L. (1990). *Predicting college grades: An analysis of institutional trends over two decades*. New York: College Entrance Examination Board and Educational Testing Service.

Wilson, K. M. (1983). *A review of research on the prediction of academic performance after the freshman year* (College Board Rep. No. 83-2). New York: College Entrance Examination Board.

Wolfe, R. N., & Johnson S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement, 55*(2), 177-185.

Woodruff, D. J., & Ziomek, R. L. (2004). *High school grade inflation from 1991 to 2003.* (ACT Research Report 04-4). Iowa City: The American College Testing Program.

Young, D. J. (1998). Rural and Urban Differences in Student Achievement in Science and Mathematics: A Multilevel Analysis. *School Effectiveness and School Improvement, 9*(4), 386-412.

Young, J, W. (1989). *Developing a universal scale for grades: Investigating predictive validity in college admissions*. Ph.D. Dissertation. Stanford University.

Young, J. W. (1991). Improving the prediction of college performance of ethnic minorities using IRT-based GPA. *Applied Measurement in Education, 4*, 229-239.

Young, J. W. (1993). Grade adjustment methods. *Review of Educational Research, 63,* 151-165.

Young, J. W. (1994). Differential prediction of college grades by gender and by ethnicity: A replicated study. *Educational and Psychological Measurement, 54,* 1022-1029.

Young, J. W. (2001). *Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis* (College Board Rep. No. 01-6).New York: College Entrance Examination Board.

Zheng, J. L., Saunders, K. P., Shelley II, M. C., & Whalen, D. F. (2002). Predictors of academic success for freshmen residence hall students. *Journal of College Student Development, 43*(2), 267-283.

Zwick, R., & Sklar, J. C. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language. *American Educational Research Journal, 42*, 439-464.